

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM protein - protein search, using SW model

Run on: April 12, 2005, 15:37:30, Search time 75.6667 Seconds

(without alignments)
636.206 Million cell updates/sec

Title: US-10-705-716A-4

Perfect score: 1 MCGGSRADAIERYRYESWT.....VTDSTQMDRSRRITKNCVN 145

Sequence:

BLOSUM62

Scoring table: Gapop 10.0, Gapext 0.5

Searched: 1418010 seqs, 33197259 residues

Total number of hits satisfying chosen parameters: 1418010

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published Applications AA:
1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubppa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubppa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubppa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubppa/US09C_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubppa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	778	100.0	145	US-10-293-239-17	Sequence 17, Appl
2	778	100.0	145	US-10-177-390-30	Sequence 30, Appl
3	750.5	96.5	180	US-10-293-239-19	Sequence 19, Appl
4	569.5	73.2	149	US-10-293-239-21	Sequence 21, Appl
5	309	39.7	73	US-10-293-239-20	Sequence 20, Appl
6	299	38.4	54	US-10-293-239-18	Sequence 18, Appl
7	293	37.7	80	US-10-293-239-12	Sequence 22, Appl
8	140	18.0	25	US-10-293-239-37	Sequence 37, Appl
9	99	12.7	18	US-10-293-239-35	Sequence 35, Appl
10	92.5	11.9	307	US-10-437-963-181279	Sequence 181279,
11	92.5	11.9	1001	US-10-415-147-3	Sequence 3, Appl
12	91	11.7	670	US-10-156-761-14107	Sequence 14107, A
13	90	11.6	219	US-10-156-761-13447	Sequence 13447, A

14	88	11.3	298	10	US-09-910-600-26	Sequence 26, Appl
15	87.5	11.2	369	14	US-10-036-547-84	Sequence 84, Appl
16	87.5	11.2	369	15	US-10-188-832-149	Sequence 149, Appl
17	87.5	11.2	369	17	US-10-658-884-4	Sequence 4, Appl
18	87.5	11.2	383	14	US-10-029-32058	Sequence 32058, A
19	87	11.2	136	15	US-10-424-599-179100	Sequence 179100,
20	86	11.1	182	16	US-10-767-701-37779	Sequence 37779, A
21	86	11.1	342	15	US-10-112-944-914	Sequence 914, Appl
22	86	11.1	668	16	US-10-437-963-134726	Sequence 134726,
23	85.5	11.0	496	15	US-10-425-114-71015	Sequence 71015, A
24	85	10.9	486	15	US-10-282-122A-62114	Sequence 62114, A
25	84	10.8	514	9	US-09-712-363-154	Sequence 154, App
26	83.5	10.7	332	16	US-10-437-963-168390	Sequence 168390,
27	83.5	10.7	558	14	US-10-156-761-12110	Sequence 12110, A
28	83	10.7	200	16	US-10-767-701-32315	Sequence 32315, A
29	83	10.7	373	16	US-10-437-963-147910	Sequence 147910,
30	83	10.7	2483	17	US-10-741-600-1170	Sequence 1170, Ap
31	82.5	10.6	497	16	US-10-437-963-157852	Sequence 157852,
32	82.5	10.6	850	15	US-10-424-599-242653	Sequence 242653,
33	82	10.5	216	16	US-10-767-701-57343	Sequence 57343, A
34	82	10.5	657	16	US-10-437-963-163549	Sequence 163549,
35	82	10.5	795	15	US-10-424-599-174901	Sequence 174901,
36	81.5	10.5	326	15	US-10-425-114-69350	Sequence 69350, A
37	81	10.4	147	16	US-10-767-701-46905	Sequence 46905, A
38	81	10.4	243	15	US-10-424-599-209325	Sequence 209325,
39	81	10.4	355	15	US-10-425-114-42733	Sequence 42733, A
40	81	10.4	2527	16	US-10-408-765A-2462	Sequence 2462, Ap
41	81	10.4	2715	16	US-10-408-765A-866	Sequence 866, App
42	80.5	10.3	213	15	US-10-424-599-190791	Sequence 190791,
43	80.5	10.3	337	16	US-10-437-963-107097	Sequence 107097,
44	80.5	10.3	342	9	US-09-269-390-4	Sequence 4, Appl
45	80.5	10.3	342	14	US-10-317-832-118	Sequence 118, Appl

ALIGNMENTS

RESULT 1	US-10-293-239-17	Application US/10293239
Sequence 17,	Application No. (US20030119043A1)	
GENERAL INFORMATION:		
APPLICANT:	Tanner, Stephan	
APPLICANT:	de la Chapelle, Albert	
TITLE OF INVENTION:	BALC expression as a diagnostic marker for acute leukemia	
FILE REFERENCE:	22727/04101	
CURRENT APPLICATION NUMBER:	US/10/293,239	
CURRENT FILING DATE:	2002-11-12	
PRIOR APPLICATION NUMBER:	US 60/346,210	
PRIOR FILING DATE:	2001-11-09	
NUMBER OF SEQ ID NOS:	39	
SOFTWARE:	PatentIn version 3.1	
SEQ ID NO 17		
LENGTH:	145	
TYPE:	PRT	
ORGANISM:	Homo sapiens	
US-10-293-239-17		
Query Match	Score 778, DB 14, Length 145;	
Best Local Similarity	100.0%; Pred. No. 1e-66;	
Matches	145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MCGGSRADAIERYRYESWTRETSTWLYTSDAPPSAAPPSPGAGLHSGMLBDGL 60	
DB	1 MCGGSRADAIERYRYESWTRETSTWLYTSDAPPSAAPPSPGAGLHSGMLBDGL 60	
QY	61 PSNGVRSRTPAGGIPNPEKKNCETCCPNPOSLSGPELTKOKGLQTTKXRPKXMPAK 120	
DB	61 PSNGVRSRTPAGGIPNPEKKNCETCCPNPOSLSGPELTKOKGLQTTKXRPKXMPAK 120	
QY	121 EVTINVTDSIQMDRSRRITKNCVN 145	
DB	121 EVTINVTDSIQMDRSRRITKNCVN 145	

RESULT 2

US-10-177-390-30
; Sequence 30, Application US/10177390
; Publication No. US20030143743A1
; GENERAL INFORMATION:
; APPLICANT: Schuler, Gerald
; APPLICANT: N.V. Antwerp Innovatiecentrum
; TITLE OF INVENTION: Improved Transfection of Eucaryotic Cells with Linear
; FILE REFERENCE: 021505wc/JH/ml
; CURRENT APPLICATION NUMBER: US/10/177,390
; CURRENT FILING DATE: 2002-06-20
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-390-30

Query Match 100.0%; Score 778; DB 14; Length 145;
Best Local Similarity 100.0%; Pred. No. 1e-66;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERPRYESWTRTESTWLTYYTDSADAPPSAAAPDSGPEAGLHSGMLEDGI 60
DB 1 MCGGSRADAIERPRYESWTRTESTWLTYYTDSADAPPSAAAPDSGPEAGLHSGMLEDGL 60

QY 61 PSNGVPRSTAPGIPNPEKKTNCETQCPNPOSLSGSLTKONGLOTTEAKRDAKMPAK 120
DB 61 PSNGVPRSTAPGIPNPEKKTNCETQCPNPOSLSGSLTKONGLOTTEAKRDAKMPAK 120

QY 121 EYTIINVTDSIQOMDRSRRTKNCVN 145
DB 121 EYTIINVTDSIQOMDRSRRTKNCVN 145

RESULT 3

US-10-293-239-19
; Sequence 19, Application US/10293239
; Publication No. US20030119043A1
; GENERAL INFORMATION:
; APPLICANT: Tanner, Stephan
; APPLICANT: de la Chapelle, Albert
; TITLE OF INVENTION: BALC expression as a diagnostic marker for acute leukemia
; FILE REFERENCE: 22727/04101
; CURRENT APPLICATION NUMBER: US/10/293,239
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/348,210
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-239-19

Query Match 96.5%; Score 750.5; DB 14; Length 180;
Best Local Similarity 80.6%; Pred. No. 6e-64;
Matches 145; Conservative 0; Mismatches 0; Indels 35; Gaps 1;

QY 1 MCGGSRADAIERPRYESWTRTESTWLTYYTDSADAPPSAAAPDSGPEAGLHSGMLEDGI 53
DB 1 MCGGSRADAIERPRYESWTRTESTWLTYYTDSADAPPSAAAPDSGPEAGLHSGMLEAERS 60

QY 54 -----GMLDGLPSNGVPRSTAPGIPNPEKKTNCET 85
DB 61 KIKAPTDVSDGLFSASKMAPLAVFSHGMLDGLPSNGVPRSTAPGIPNPEKKTNCET 120

QY 86 QCPNPOSLSGSLTKONGLOTTEAKRDAKMPAK 145

DB

121 QCPNPOSLSGSLTKONGLOTTEAKRDAKMPAK 180

RESULT 4

US-10-293-239-21
; Sequence 21, Application US/10293239
; Publication No. US20030119043A1
; GENERAL INFORMATION:
; APPLICANT: Tanner, Stephan
; APPLICANT: de la Chapelle, Albert
; TITLE OF INVENTION: BALC expression as a diagnostic marker for acute leukemia
; FILE REFERENCE: 22727/04101
; CURRENT APPLICATION NUMBER: US/10/293,239
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/348,210
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-239-21

Query Match 73.2%; Score 569.5; DB 14; Length 149;
Best Local Similarity 75.7%; Pred. No. 1.1e-46;
Matches 109; Conservative 0; Mismatches 0; Indels 35; Gaps 1;

QY 1 MCGGSRADAIERPRYESWTRTESTWLTYYTDSADAPPSAAAPDSGPEAGLHSGMLEDGI 53
DB 1 MCGGSRADAIERPRYESWTRTESTWLTYYTDSADAPPSAAAPDSGPEAGLHSGMLEAERS 60

QY 54 -----GMLDGLPSNGVPRSTAPGIPNPEKKTNCET 85
DB 61 KIKAPTDVSDGLFSASKMAPLAVFSHGMLDGLPSNGVPRSTAPGIPNPEKKTNCET 120

QY 86 QCPNPOSLSGSLTKONGLOTTE 109
DB 121 QCPNPOSLSGSLTKONGLOTTE 144

RESULT 5

US-10-293-239-20
; Sequence 20, Application US/10293239
; Publication No. US20030119043A1
; GENERAL INFORMATION:
; APPLICANT: Tanner, Stephan
; APPLICANT: de la Chapelle, Albert
; TITLE OF INVENTION: BALC expression as a diagnostic marker for acute leukemia
; FILE REFERENCE: 22727/04101
; CURRENT APPLICATION NUMBER: US/10/293,239
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/348,210
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-239-20

Query Match 39.7%; Score 309; DB 14; Length 73;
Best Local Similarity 96.6%; Pred. No. 4.6e-22;
Matches 56; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERPRYESWTRTESTWLTYYTDSADAPPSAAAPDSGPEAGLHSGMLEDGI 58
DB 1 MCGGSRADAIERPRYESWTRTESTWLTYYTDSADAPPSAAAPDSGPEAGLHSGLCLE 58

RESULT 6

```

US-10-293-239-18
; Sequence 18, Application US/10293239
; Publication No. US20030119043A1
; GENERAL INFORMATION:
; APPLICANT: Tanner, Stephan
; APPLICANT: de la Chapelle, Albert
; FILE REFERENCE: 22727/04101
; CURRENT APPLICATION NUMBER: US/10/293,239
; PRIOR FILING DATE: 2002-11-12
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 18
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-239-18

Query Match
Best Local Similarity 38.4%; Score 299; DB 14; Length 54;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy
1 MCGGSRADAIERPRYESWTRTESTWLTYYTDSADAPPSPAAPDSGPEAGGLHSG 54
Db
1 MCGGSRADAIERPRYESWTRTESTWLTYYTDSADAPPSPAAPDSGPEAGGLHSG 54

RESULT 7
US-10-293-239-22
; Sequence 22, Application US/10293239
; Publication No. US20030119043A1
; GENERAL INFORMATION:
; APPLICANT: Tanner, Stephan
; APPLICANT: de la Chapelle, Albert
; FILE REFERENCE: 22727/04101
; CURRENT APPLICATION NUMBER: US/10/293,239
; PRIOR FILING DATE: 2002-11-12
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 22
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-239-22

Query Match
Best Local Similarity 37.7%; Score 293; DB 14; Length 80;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy
1 MCGGSRADAIERPRYESWTRTESTWLTYYTDSADAPPSPAAPDSGPEAGGLHSG 53
Db
1 MCGGSRADAIERPRYESWTRTESTWLTYYTDSADAPPSPAAPDSGPEAGGLHSG 53

RESULT 8
US-10-293-239-37
; Sequence 37, Application US/10293239
; Publication No. US20030119043A1
; GENERAL INFORMATION:
; APPLICANT: Tanner, Stephan
; APPLICANT: de la Chapelle, Albert
; FILE REFERENCE: 22727/04101
; CURRENT APPLICATION NUMBER: US/10/293,239
; PRIOR FILING DATE: 2002-11-12
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 37
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-239-37

```

```

; SOFTWARE: Patentin version 3.1
; SEQ ID NO 37
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-239-37

Query Match
Best Local Similarity 18.0%; Score 140; DB 14; Length 25;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy
7 RADAIERPRYESWTRTESTWLTYYT 31
Db
1 RADAIERPRYESWTRTESTWLTYYT 25

RESULT 9
US-10-293-239-35
; Sequence 35, Application US/10293239
; Publication No. US20030119043A1
; GENERAL INFORMATION:
; APPLICANT: Tanner, Stephan
; APPLICANT: de la Chapelle, Albert
; FILE REFERENCE: 22727/04101
; CURRENT APPLICATION NUMBER: US/10/293,239
; PRIOR FILING DATE: 2002-11-12
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 35
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-239-35

Query Match
Best Local Similarity 12.7%; Score 99; DB 14; Length 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy
9 DAIEPRYESWTRTEST 26
Db
1 DAIEPRYESWTRTEST 18

RESULT 10
US-10-437-963-181279
; Sequence 181279, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; PRIOR FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 181279
; LENGTH: 307
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(307)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-437-963-181279

```

[illegible]

```

RESULT 11
US-10-415-147-3
; Sequence 3, Application US/10415147
; Publication No. US20040043399A1
; GENERAL INFORMATION:
; APPLICANT: Bayer AG
; TITLE OF INVENTION: REGULATION OF HUMAN INOSITOL POLYPHOSPHATE 5- PHOSPHATASE
; FILE REFERENCE: L10235foreign Countrie$
; CURRENT APPLICATION NUMBER: US/10/415,147
; CURRENT FILING DATE: 2003-04-25
; PRIOR APPLICATION NUMBER: US 60/243,745
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: US 60/257,302
; PRIOR FILING DATE: 2000-12-26
; PRIOR APPLICATION NUMBER: US 60/314,660
; PRIOR FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1001
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; US-10-415-147-3

```

	Query Match	Similarity	Score	IDB	Length	Mismatches	Conservative	Indels	Gaps
		11.9%	92.5	DB 15	1001				
	Best Local Similarity	25.8%	Pred. No. 7						
	Matches	32	Conservative	15	Mismatches	48	Indels	29	Gaps
QY	12 EPRYESWTRETSTFWLTYTDSAP	-----SAAAPDSCGEAGLHSGLTEDGLPS	62						
Db	275 DRLSPFARAEARRHSBEDVLP	PPRQTLPLDVSSGLPESGRSRGSLPFRRGIRS	334						
QY	63 NGVPNSTABGILPNFEK	-----KTNC-----ETOCENFGSISSGPLTKONGLO	106						
Db	335 N-----QTVPPLPKPPRSRSPSRKS	PMRSPCVPAPEVALPRVTGTGAAGFKCSPENIQ	390						
QY	107 TTEA 110								
Db	391 TQES 394								

RESULT 12
 US-10-156-761-14107
 ; Sequence 14107, Application US/10156761
 ; Publication No. US20030112018A1
 GENERAL INFORMATION
 ;
 ; APPLICANT: OMURA, SATOSHI
 ; APPLICANT: IKEDA, HARUO
 ; APPLICANT: ISHIKAWA, JUN
 ; APPLICANT: HORIKAWA, HIROSHI
 ; APPLICANT: SHIBA, TADAOSHI
 ; APPLICANT: SAKAKI, YOSHITUKU

```

1  APPLICANT: HATTORI, MASAHIRA
2  TITLE OR INVENTION: 'NOVEL POLYNUCLEOTIDES
3  FILE REFERENCE: 249-252
4  CURRENT APPLICATION NUMBER: US/10/156,761
5  CURRENT FILING DATE: 2002-05-29
6  PRIOR APPLICATION NUMBER: JP 2001-204089
7  PRIOR FILING DATE: 2001-05-30
8  PRIOR APPLICATION NUMBER: JP 2001-272697
9  PRIOR FILING DATE: 2001-08-02
10 NUMBER OF SEQ ID NOS: 15109
11 SEQ ID NO 14107
12 LENGTH: 670
13 TYPE: PRN
14 ORGANISM: Streptomyces avermitilis
15 US-10-156-761-14107

```

Query Match	11.7%;	Score 91;	DB 14;	Length 670;
Best Local Similarity	25.4%;	Pred. No. 6;		
Matches 32;	Conservative 14;	Mismatches 50;	Indels 30;	Gaps 4

QY 5 GSRDAIERYEYESTRETETSTWLTYYDDAPPS--AAAPDSGEPEAGLHSGWLEBGLPS 62

Db 16 GHRGEGMSPR-----TNDAPESDGEOPPAPEAGTAETGPASATGT 59

QY 63 NGVPNSTAPGGIIPNEKKTNCETOCNPQSLSSGPILOXONGQTTE-----AKRDAX 115

Db 60 AGAPMEESAGGATTEEBPVGATTEKP---SAGATTEDEPSAGATTEESPADAVAOQSSAA 114

QY 116 RMPAKE 121

Db 115 AAPAE 120

```

RESULT 13
US-10-156-761-13447
# Sequence 13447, Application US/10156761
# Publication No. US20030119018A1
# GENERAL INFORMATION:
# APPLICANT: OMURA, SATOSHI
# APPLICANT: IKEDA, HARUO
# APPLICANT: ISHIKAWA, JUN
# APPLICANT: HOSIKAWA, JUN
# APPLICANT: SHIBA, TADAYOSHI
# APPLICANT: SAWAKI, YOSHIYUKI
# APPLICANT: HATTORI, MASAHIRA
# TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
# FILE REFERENCE: 249-262
# CURRENT APPLICATION NUMBER: US/10-156,761
# CURRENT FILING DATE: 2002-05-29
# PRIOR APPLICATION NUMBER: JP 2001-204089
# PRIOR FILING DATE: 2001-05-30
# PRIOR APPLICATION NUMBER: JP 2001-272697
# PRIOR FILING DATE: 2001-08-02
# NUMBER OF SEQ ID NOS: 15109
# SEQ ID NO 13447
# LENGTH: 219
# TYPE: PRT
# ORGANISM: Streptomyces avermitilis
# US-10-156-761-13447

```

Query Match: 11.6%; Score: 90; DB: 14; Length: 219;
Best Local Similarity: 33.9%; Pred. No.: 1,9;
Matches: 20; Conservative: 8; Mismatches: 31; Indels: 0; Gaps: 0;

QY 6 SBAADIERRYESTRETETSTWLTYYDDSDAPPSAALPDGSPRAGIHSMTLEDGLPNSG 64
AAAGVGGKOTITRWMSASAEVILBAFTDLSQAEEAARBPPEGGOGGGNGNGLPDIG 101

RESULT 14
US-09-910-600-26
; Sequence 26, Application US/09910600
; Publication No. US20030036631A1

This Page Blank (uspto)

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension .rup) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

This Page Blank (uspto)

CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention

CC Sequence 145 AA;

Query Match 100.0%; Score 778; DB 4; Length 145;
 Best Local Similarity 100.0%; Pred. No. 4e-70; Mismatches 0; Gaps 0;
 Matches 145; Conservative 0; Indels 0; Gaps 0;

QY 1 MCGGSRADAIIPRYESWTRETSTWLTYSDDAPPSAAAPDSGEGAGLHSGMLDGL 60
 DB 1 MCGGSRADAIIPRYESWTRETSTWLTYSDDAPPSAAAPDSGEGAGLHSGMLDGL 60
 QY 61 PANGVPRSTAPGGINPEKKTNCETCCPPOSLSGGLTQKONGLOTTEAKDKAKMPAK 120
 DB 61 PANGVPRSTAPGGINPEKKTNCETCCPPOSLSGGLTQKONGLOTTEAKDKAKMPAK 120
 QY 121 EVTINVTDSIQMDRSRRITKNCVN 145
 DB 121 EVTINVTDSIQMDRSRRITKNCVN 145

RESULT 2

AA019498
 ID AA019498 standard; protein; 145 AA.

AC AA019498;

DT 20-DEC-2002 (first entry)

DE Hs1 protein variant.

XX Hs1; variant; cancer; tumour; unigene cluster; cytostatic; metastasis;
 KW E87; expressed sequence tag; colon cancer; stomach cancer; breast cancer;
 KW HS169395; HS127144; HS2; HS132793; HS3.

XX Unidentified.

XX DE10103694-A1.

XX 26-JAN-2001; 2001DE-01003694.

XX 26-JAN-2001; 2001DE-01003694.

XX 26-JAN-2001; 2001DE-01003694.

XX (DELB-) DELBRUECK CENT MOLEKULARE MEDIZIN MAX.

XX Bietz, Kemmer W;

XX WPI; 2002-644836/70.

XX N-PSDB; AAL50100.

PT Diagnosis and therapy of tumors, by determining expression rates of
 PT specific expressed sequence tags of the unigene cluster, and subsequently
 PT blocking their expression.

PS Claim 10; Page 5; 10pp; German.

CC The present invention relates to the use of expressed sequence tags
 (ESTs), or variants, of the unigene cluster HS169395 (HS1), HS127144

CC (HS2) and/or HS132793 (HS3) for diagnosis and therapy of tumors, in
 CC which their expression rates in tumour cells and/or lymph nodes are
 CC determined. The EST sequences are useful as prognostic markers of
 CC survival of cancer patients (high levels of EST-related mRNA are
 CC associated with a poor prognosis, specifically correlated with
 CC development of metastases); and for diagnosis and/or therapy of solid
 CC tumours, particularly of colon, stomach and breast. The present sequence
 CC is a variant of the Hs1 protein shown in the exemplification of the
 CC invention

CC Sequence 145 AA;

Query Match 100.0%; Score 778; DB 5; Length 145;
 Best Local Similarity 100.0%; Pred. No. 4e-70; Mismatches 0; Gaps 0;
 Matches 145; Conservative 0; Indels 0; Gaps 0;

QY 1 MCGGSRADAIIPRYESWTRETSTWLTYSDDAPPSAAAPDSGEGAGLHSGMLDGL 60
 DB 1 MCGGSRADAIIPRYESWTRETSTWLTYSDDAPPSAAAPDSGEGAGLHSGMLDGL 60
 QY 61 PANGVPRSTAPGGINPEKKTNCETCCPPOSLSGGLTQKONGLOTTEAKDKAKMPAK 120
 DB 61 PANGVPRSTAPGGINPEKKTNCETCCPPOSLSGGLTQKONGLOTTEAKDKAKMPAK 120
 QY 121 EVTINVTDSIQMDRSRRITKNCVN 145
 DB 121 EVTINVTDSIQMDRSRRITKNCVN 145

RESULT 3

ABR58646
 ID ABR58646 standard; protein; 145 AA.

AC ABR58646;

DT 09-JUL-2003 (first entry)

DE Human cancer related protein SEQ ID NO:303.

XX Human; cancer; diagnosis; screening; modulator; leukaemia; leaemia;

XX heart disease; atherosclerosis; endometriosia.

XX Homo sapiens.

XX WO2003025138-A2.

XX 27-MAR-2003.

XX 17-SEP-2002; 2002WO-US029560.

XX 17-SEP-2001; 2001US-0323469P.

XX 20-SEP-2001; 2001US-0323887P.

XX 13-NOV-2001; 2001US-0350666P.

XX 08-FEB-2002; 2002US-035145P.

XX 08-FEB-2002; 2002US-0355257P.

XX 12-APR-2002; 2002US-0372246P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Afar D, Aziz N, Gish KC, Hevezi PA, Mack DH, Wilson KE,

XX Zlocnik A;

XX WPI; 2003-354600/33.

XX N-PSDB; ACC72796.

PT New genes that are up-regulated or down-regulated in cancers, useful as
 PT markers for diagnosing e.g. cancer, ischemia or heart diseases, or as
 PT therapeutic targets for screening drugs for treating these diseases.

PS Claim 12; Page 753; 767pp; English.

CC The present invention describes an isolated nucleic acid molecule, which
 CC comprises the sequence of any of the genes that are up-regulated or down-

1.131 de c lumbis

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 15:15:13 ; Search time 103.667 Seconds
(without alignment)

540.967 Million cell updates/sec

Title: US-10-705-716a-4

Perfect score: 778
Sequence: 1 MCGCGSRRAIRRYESWT.....VDSIQMDRSRRITKNCVN 145

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	778	100.0	145	4	AA895018 Aab95018 Human pro
2	778	100.0	145	5	AA019498 Aa019498 HSI prote
3	778	100.0	145	6	ABR58646 Abt58646 Human can
4	778	100.0	145	7	AD031800 Adc31800 Human nov
5	778	100.0	145	8	ADMA6959 Adma6959 Brain and
6	778	100.0	145	8	ADMA6959 Adma6959 Human PTH
7	750.5	96.5	180	7	ADMA6961 Adma6961 Brain and
8	649	83.4	145	8	ADMA6961 Adma6961 Mouse PTH
9	645	82.9	145	8	ADMA6961 Adma6961 Rat PTH
10	569.5	73.2	149	7	ADMA6963 Adma6963 Brain and
11	405.5	52.1	92	5	ADQ81902 Adq81902 Human dlo
12	309	39.7	73	7	ADMA6962 Adma6962 Brain and
13	299	38.4	54	7	ADMA6960 Adma6960 Brain and
14	299	38.4	54	8	ADMA6960 Adma6960 Human PTH
15	293	37.7	54	8	ADMA6964 Adma6964 Brain and
16	278	35.7	54	8	ADMA6964 Adma6964 Mouse PTH
17	140	18.0	25	7	ADMA6979 Adma6979 Brain and
18	99	12.7	18	7	ADMA6979 Adma6979 Brain and
19	92.5	11.9	1001	5	AAU98903 Aau98903 Rat inos
20	91.5	11.8	210	3	AAU98903 Aau98903 Arabidops
21	91.5	11.8	222	3	AAU98903 Aau98903 Arabidops
22	91.5	11.8	222	3	AAU98903 Aau98903 Arabidops
23	89	11.4	16	8	ADMA6979 Adma6979 Brain and
24	89	11.4	321	4	AB871629 Abb871629 Drosophi
25	88	11.3	298	5	AAU87066 Aau87066 Sg1lec-BM

26	87.5	11.2	369	3	AAU71485 Aay71485 Human MAG
27	87.5	11.2	369	4	AA800297 Aab800297 Human pro
28	87.5	11.2	369	6	ABR48215 Abt48215 Human bla
29	87.5	11.2	369	6	ABU56516 Abu56516 Lung canc
30	87.5	11.2	383	8	ABO58424 Abo58424 Human gen
31	86.5	11.1	1192	4	ABG02038 Abg02038 Novel hum
32	86	11.1	342	8	ADM87821 Adm87821 Human EST
33	86	11.1	508	4	ABG09910 Abg09910 Novel hum
34	86	11.1	586	5	ABD04711 Abd04711 Human PPI
35	86	11.1	991	4	ABR83195 Aab83195 Human Rec
36	86	11.1	991	4	ADO55153 Ado55153 Protein #
37	85.5	11.0	211	8	ABM81408 Abm81408 Tumour-as
38	85.5	11.0	236	8	AD510476 Ad510476 Human the
39	85	10.9	486	6	ABU34190 Abu34190 Protein e
40	84.5	10.9	1111	5	AAU71108 Aau71108 Murine G1
41	84	10.8	514	4	AA681103 Aa681103 Mycobacte
42	83.5	10.7	684	4	ABR69330 Abb69330 Drosophi
43	83.5	10.7	899	4	ABR65489 Abb65489 Drosophi
44	83.5	10.7	899	4	ABR65489 Abb65489 Drosophi
45	83	10.7	659	4	ABR65280 Abb65280 Drosophi

ALIGNMENTS

RESULT-1
ID AAB95018 standard; protein; 145 AA.
XX
AC AAB95018;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:16726.
XX
KM Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
XX
PR 27-AUG-1999; 99JP-00300253.
XX
PR 11-JAN-2000; 2000JP-00118776.
XX
PR 02-MAY-2000; 2000JP-00183767.
XX
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
XX
PT length cDNAs defined in the specification, and for the detection and/or
XX
PT diagnosis of the abnormality of the proteins encoded by the full-length
XX
PT cDNAs.
XX
PS Claim 8; SEQ ID NO 16726; 2537pp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesizing 5602 full-
XX
CC length cDNAs defined in the specification, where a primer set comprises:
XX
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
XX
CC complementary strand of a polynucleotide which comprises one of the 5602
XX
CC nucleotide sequences defined in the specification, where the
XX
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX
CC of an oligonucleotide comprising a sequence complementary to the
XX
CC complementary strand of a polynucleotide which comprises a 5'-end

CC	sequence and an oligonucleotide comprising a sequence complementary to a
CC	polynucleotide which comprises a 3'-end sequence, where the
CC	oligonucleotide comprises at least 15 nucleotides and the combination of
CC	the 5'-end sequence/3'-end sequence is selected from those defined in the
CC	specification. The primer sets can be used in antisense therapy and in
CC	gene therapy. The primers are useful for synthesizing polynucleotides,
CC	particularly full-length cDNAs. The primers are also useful for the
CC	detection and/or diagnosis of the abnormality of the proteins encoded by
CC	the full-length cDNAs. The primers allow obtaining of the full-length
CC	cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC	AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC	represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC	oligonucleotides, all of which are used in the exemplification of the
CC	present invention
XX	
XX	Sequence 145 AA;
XX	
XX	Query Match 100.0%; Score 778; DB 4; Length 145;
XX	Best Local Similarity 100.0%; Pred. No. 4e-70;
XX	Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY	1 MCGGGSRADAIEPRYYESWTRETESTWLTYYTDSAPPSAAA PDGPEAGLHSGMLDGL 60
DB	1 MCGGGSRADAIEPRYYESWTRETESTWLTYYTDSAPPSAAA PDGPEAGLHSGMLDGL 60
QY	61 PSNGVFRSTAPGCIIPPEKKTNCFCQCPPOSLSGCLPTQKNGLOTTTEAKRDAKMPAK 120
DB	61 PSNGVFRSTAPGCIIPPEKKTNCFCQCPPOSLSGCLPTQKNGLOTTTEAKRDAKMPAK 120
QY	121 EVTINVTDSIQQMDRSRRITTKACVN 145
DB	121 EVTINVTDSIQQMDRSRRITTKACVN 145
XX	
XX	RESULT 2
XX	AA019498
XX	ID AA019498 standard; protein; 145 AA.
XX	AA019498;
XX	20-DEC-2002 (first entry)
XX	20-DEC-2002 (first entry)
XX	HS1 protein variant.
XX	HS1; variant; cancer; tumour; uijene cluster; cytostatic; metastasis;
XX	KW EST; expressed sequence tag; colon cancer; stomach cancer; breast cancer;
XX	KM HS169395; HS127144; HS2; HS132793; HS3.
XX	Unidentified.
XX	OS
XX	FN DE10103694-AA1.
XX	PD 01-AUG-2002.
XX	PF 26-JAN-2001; 2001DE-01003694.
XX	PR 26-JAN-2001; 2001DE-01003694.
XX	PA (DELBA-) DELBRUECK CENT MOLEKULARE MEDIZIN MAX.
XX	PI Brett D, Kemmerer W;
XX	DR WPI; 2002-644836/70.
XX	DR N-PSDB; AAL50100.
XX	diagnosis and therapy of tumors, by determining expression rates of
XX	specific expressed sequence tags of the ungine cluster, and subsequently
XX	blocking their expression.
XX	Claim 10; Page 5; 10pp; German.
XX	The present invention relates to the use of expressed sequence tags
XX	(ESTs), or variants, of the ungine cluster HS169395 (HS1), HS127144

(HS2) and/or HS137793 (HS3) for diagnosis and therapy of tumours, in which their expression rates in tumour cells and/or lymph nodes are determined. The EST sequences are useful as prognostic markers of survival of cancer patients (high levels of EST-related mRNA are associated with a poor prognosis, specifically correlated with development of metastases); and for diagnosis and/or therapy of solid tumours, particularly of colon, stomach and breast. The present sequence is a variant of the HSI protein shown in the exemplification of the invention

Sequence 145 AA;

Query Match 100.0%; Score 778; DB 5; Length 145;
Best Local Similarity 100.0%; Pred. No. 4e-70;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Dc
1 MCGGSRADAIIEPRYSESWTEETESTWLTLYTDSAPSAAPDGGPAGGLHSGMLEDL 60
1 MCGGSRADAIIEPRYSESWTRTESTWLTLYTDSAPSAAPDGGPAGGLHSGMLEDL 60

Dd
61 PSNGVPRSTAPGGINPEKTKNCTQCNPQSLSSGGPLTKONGLOTTEAKRDARMPAK 120
61 PSNGVPRSTAPGGINPEKTKNCTQCNPQSLSSGGPLTKONGLOTTEAKRDARMPAK 120

Dy
121 EVTINTDSIQMDRSRRITKNCVN 145
121 EVTINTDSIQMDRSRRITKNCVN 145

Dz
121 EVTINTDSIQMDRSRRITKNCVN 145
121 EVTINTDSIQMDRSRRITKNCVN 145

RESULT 3
ABR58646
ID ABR58646 standard; protein; 145 AA.
XX
AC ABR58646;
XX
DT 09-JUL-2003 (first entry)
XX
DE Human cancer related protein SEQ ID NO:303.
XX
KW Human; cancer; diagnosis; screening; modulator; leukaemia; ischaemia;
KM heart disease; atherosclerosis; endometriosis.
XX
OS Homo sapiens.
XX
PN WC0003025138-AA.
XX
PP 27-MAR-2003.
XX
PF 17-SEP-2002; 2002WO-US029560.
XX
PR 17-SEP-2001; 2001US-0323469P.
PR 20-SEP-2001; 2001US-032387P.
PR 13-NOV-2001; 2001US-035066P.
PR 08-FEB-2002; 2002US-0355145P.
PR 08-FEB-2002; 2002US-0355257P.
PR 12-APR-2002; 2002US-0372246P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
PI Ahar D, Aziz N, Gish KC, Hevezl PA, Mack DH, Wilson KE;
FI Zlotnick A;
XX
DR WPI; 2003-354600/33.
XX
DR N-PADB; ACCT2796.
XX
PT New genes that are up-regulated or down-regulated in cancers, useful as
CC markers for diagnosing e.g. cancer, ischemia or heart diseases, or as
XX therapeutic targets for screening drugs for treating these diseases.
PS Claim 12; Page 753; 767pp; English.

The present invention describes an isolated nucleic acid molecule, which comprises the sequence of any of the genes that are up-regulated or down-

1.13/ de c barkin

1.13/dec bank

CC regulated in specific cancers (e.g. about 1031 genes up-regulated in
CC acute lymphocytic leukemia). ACC72641 to ACC72860 represent cancer
CC related gene nucleotide sequences which encode the proteins given in
CC ABR58521 to ABR58709. Also described: (1) determining the presence or
CC absence of a pathological cell in a patient; (2) an expression vector
CC comprising a nucleic acid molecule described above; (3) a host cell
CC comprising the vector; (4) an isolated polypeptide, which is encoded by
CC the nucleic acid; (5) an antibody that specifically binds the polypeptide
CC of (4); (6) specifically targeting a compound to a pathological cell in a
CC patient by administering to the patient the antibody above; and (7) a
CC drug screening assay. The nucleic acid is useful as diagnostic markers or
CC therapeutic targets. In particular, the nucleic acid is useful for
CC diagnosing a pathology, e.g. cancer (e.g. cancer of the bone marrow,
CC bladder, brain, breast, cervix, colon/rectum, kidney, lung, ovary,
CC pancreas, prostate, skin and uterus), wounds, ischaemia, heart diseases,
CC atherosclerosis and endometriosis. The nucleic acid is also useful in
CC drug screening, particularly for identifying agents for treating these
CC pathologies
XX
SQ Sequence 145 AA:

Query Match 100.0%; Score 778; DB 6; Length 145;
Best Local Similarity 100.0%; Pred. No. 4e-70;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERPRYESTRETSTWLTYYTSDAPPSAAPPSPGEGAGLSGMEEDGL 60
|||||
DB 1 MCGGSRADAIERPRYESTRETSTWLTYYTSDAPPSAAPPSPGEGAGLSGMEEDGL 60
|||||
QY 61 PSNGVPRSTAPGIPPEKKTNCETCCPNPSSSGPLTQKONGLOTTAKRDAKMPAK 120
|||||
DB 61 PSNGVPRSTAPGIPPEKKTNCETCCPNPSSSGPLTQKONGLOTTAKRDAKMPAK 120
|||||
QY 121 EVTINVTDSIQMDRSRRITKNCVN 145
|||||
DB 121 EVTINVTDSIQMDRSRRITKNCVN 145
|||||

RESULT 4
ADCC1800
ID ADCC1800 standard; protein; 145 AA.
XX
AC ADCC1800;
XX

DT 18-DEC-2003 (first entry)
XX

DE Human novel polypeptide sequence, SEQ ID NO:1882.
XX

KM Human: diagnostic; drug screening; forensics; gene mapping;
KM biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KM neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KM ulcers; osteoporosis; autoimmune disease; cancer;
KM molecular weight marker; food supplement; antiparkinsonian; nootropic;
KM neuroprotective; antianemic; anticoagulant; thrombolytic; vulnery;
KM anticancer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KM gene therapy; chromosome 8.
XX

OS Homo sapiens.
XX
PN WO2003029271-A2.
XX
PD 10-APR-2003.
XX
PF 24-SEP-2002; 2002WO-US030474.
XX
PR 24-SEP-2001; 2001US-0324631P.
XX
PA (HYBR-) HYSEQ INC.
XX

PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T,
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Wang G,
PI Haley-Vicente D, Drmanac RT;
XX

DR WPI; 2003-371961/35.
DR N-PSDB; ADCC08829.
XX

PT New polynucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anaemias, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
XX cancer.

PS Claim 20; SEQ ID NO 1882; 1185bp; English.

XX
XX The invention relates to 971 novel human cDNA sequences (ADCC39919-
CC ADCC0889) and the polypeptides they encode (ADCC0890-ADCC1860). The
CC invention also relates to nucleic acid sequences over 99% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; an antibody
CC against a polypeptide of the invention; a method of detecting
CC polynucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The
CC invention further discloses methods of preventing, treating or
CC ameliorating a medical condition; kits comprising polynucleotide probes
CC and/or monoclonal antibodies for carrying out the methods of the
CC invention; methods for the identification of compounds that modulate the
CC expression or activity of the polynucleotide and/or polypeptide; and 767
CC contig sequences corresponding to the cDNA sequences of the invention
CC (ADCC1861-ADCC2627) and the polypeptides encoded by the contigs (ADCC2628
CC -ADCC3394). The nucleic acids and polypeptides of the invention are
CC useful in diagnostics, drug screening, forensics, gene mapping, in the
CC identification of mutations responsible for genetic disorders or other
CC traits, for assessing biodiversity, and in producing many other types of
CC data and products dependent on DNA and amino acid sequences. They are
CC also used for treating diseases such as Parkinson's disease, Alzheimer's
CC disease and other neurodegenerative diseases, anaemia, platelet
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC cancer. The nucleic acids may also be used as hybridisation probes or
CC primers, and in the recombinant production of a protein. The polypeptides
CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a specifically
CC claimed human polypeptide sequence of the invention. Note: The sequence
CC data for this patent did not form part of the printed specification, but,
CC was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 145 AA:

Query Match 100.0%; Score 778; DB 7; Length 145;
Best Local Similarity 100.0%; Pred. No. 4e-70;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERPRYESTRETSTWLTYYTSDAPPSAAPPSPGEGAGLSGMEEDGL 60
|||||
DB 1 MCGGSRADAIERPRYESTRETSTWLTYYTSDAPPSAAPPSPGEGAGLSGMEEDGL 60
|||||
QY 61 PSNGVPRSTAPGIPPEKKTNCETCCPNPSSSGPLTQKONGLOTTAKRDAKMPAK 120
|||||
DB 61 PSNGVPRSTAPGIPPEKKTNCETCCPNPSSSGPLTQKONGLOTTAKRDAKMPAK 120
|||||
QY 121 EVTINVTDSIQMDRSRRITKNCVN 145
|||||
DB 121 EVTINVTDSIQMDRSRRITKNCVN 145
|||||

RESULT 5
ADMA46959
ID ADMA46959 standard; protein; 145 AA.
XX
AC ADMA46959;
XX

DT 03-JUN-2004 (first entry)
XX

DE Brain and Acute Leukemia, Cytoplasmic alternate protein #1.
XX

KM acute myelogenous leukemia; gene expression; BALC;
XX

KW chronic myelogenous leukemia; prostate cancer; Brain and Acute Leukemia;
KM Cytoplasmic; exon.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT MISC-difference 41
PT /note= "encoded by GCS"
XX
PN WO2003040347-A2.
XX
PD 15-MAY-2003.
XX
PF 12-NOV-2002; 2002WO-US036375.
XX
PR 09-NOV-2001; 2001US-0348210P.
XX
PA (OHIS) UNITV OHIO STATE RES FOUND.
XX
PI Tanner SM, De la Chapell A;
XX
DR MPI: 2003-441564/41.
DR N-PSDB; ADM46951.
XX
PT Characterizing acute or chronic myelogenous leukemia, or prostate cancer
PT in a patient comprises assaying for the overexpression of one or more
PT BAALC transcripts in cells obtained from the patient.
XX
PS Disclosure; SEQ ID NO 17; 78pp; English.
XX
CC The invention relates to a method of characterizing acute myelogenous
CC leukemia (AML) in a patient by assaying for the overexpression of one or
CC more BAALC transcripts in cells obtained from the patient, where an
CC overexpression indicates that the patient has an aggressive form of AML.
CC The methods, kits and probes are useful for characterizing acute or
CC chronic myelogenous leukemia, or prostate cancer. They are also useful
CC for detecting BAALC overexpression. This sequence corresponds to a BAALC
CC (Brain and Acute Leukemia, Cytoplasmic) protein encoded by an alternative
CC spliced RNA consisting of exons 1, 6 and 8.
XX
SQ Sequence 145 AA;
Query Match 100.0%; Score 778; DB 7; Length 145;
Best Local Similarity 100.0%; Pred. No. 4e-70;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MCCGSGRADAIIEPRRYESWTRTESWLTYYTDSADAPSAAPDSGPEAGLHSGMLEDEL 60
DB 1 MCCGSGRADAIIEPRRYESWTRTESWLTYYTDSADAPSAAPDSGPEAGLHSGMLEDEL 60
QY 61 PSNGVPRSTAPGCIIPNPEKKNCECTCCPNPSSISGPLETKONGIQTTEAKRDARMPAK 120
DB 61 PSNGVPRSTAPGCIIPNPEKKNCECTCCPNPSSISGPLETKONGIQTTEAKRDARMPAK 120
QY 121 EVTINVTDSIQMDRSRRITKNCVN 145
DB 121 EVTINVTDSIQMDRSRRITKNCVN 145
RESULT 6
ADO48475
ID ADO48475 standard; protein; 145 AA.
XX
AC ADO48475;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human PTH responsive gene protein.
XX
XX PTH responsive gene; PARGB; bone-forming; bone; bone density modulation;
KM transgenic animal; osteopathic; gene therapy; osteoporosis; human.
XX
OS Homo sapiens.

XX
PN WO2004044152-A2.
XX
PD 27-MAY-2004.
XX
PF 10-NOV-2003; 2003WO-US035655.
XX
PR 12-NOV-2002; 2002US-0425532P.
XX
PA (AMRP) WYETH.
XX
PI Robinson JA, Stojanovic-Susanlic V, Babić P, Murrills RJ;
XX
DR MPI: 2004-420299/39.
DR N-PSDB; ADO48474.
XX
PT New nucleic acid fragment encoding a PARGB polypeptide, useful in
PT preparing a composition for diagnosing, treating or preventing bone
PT related disorders, e.g., osteoporosis.
XX
PS Claim 9; SEQ ID NO 4; 16pp; English.
XX
CC The invention relates to a novel PTH responsive gene (PARGB) fragment
CC encoding a polypeptide. The invention further comprises: a chimeric
CC construct comprising the isolated nucleic acid fragment operatively
CC linked to suitable regulatory sequences; a host cell transformed with the
CC chimeric construct; a vector comprising the nucleic acid fragment;
CC obtaining a nucleic acid fragment encoding the polypeptide; a method for
CC obtaining a polypeptide; detecting the presence of the nucleic acid
CC fragment; an antibody that specifically binds to one or more epitopes of
CC a PARGB polypeptide; a composition for regulating bone-forming activity
CC in a mammal comprising the nucleic acid fragment, polypeptide or antibody
CC ; an agent that alters the expression of PARGB gene or polypeptide;
CC determining whether an agent alters the expression of PARGB mRNA;
CC screening agents for effectiveness in altering expression of the nucleic
CC acid fragment; screening for agents useful for treating bone related
CC disorders; evaluating the efficacy of a treatment of a bone related
CC disorder in a subject; identifying polypeptides capable of binding to
CC PARGB; monitoring the effectiveness of treatment of a subject with a bone
CC related agent; a transgenic animal comprising the DNA; an animal model
CC for the study of bone density modulation comprising a first group of
CC animals composed of the transgenic animal and a second group of control
CC animals; studying bone mass determinants; studying the modulation of bone
CC mass; studying an effect of PARGB on bone disorders; identifying an agent
CC for treating bone related disorders; identifying whether an agent which
CC has bone forming activity; and a stably transfected cell line comprising
CC two constructs, the first construct comprising a ligand binding domain
CC linked to a DNA binding domain which is linked to an activation domain
CC all of which expression is driven by a constitutive promoter, the second
CC construct comprising multiple copies of DNA binding elements linked to a
CC minimal promoter which is linked to PARGB cDNA, where upon the addition
CC of chemical inducer, transcription of PARGB gene is induced. The PARGB
CC polynucleotide has osteopathic activity. The PTH responsive gene may be
CC used to treat disorders by gene therapy. The nucleic acid is useful in
CC preparing a composition for diagnosing, treating or preventing bone
CC related disorders, e.g., osteoporosis. This sequence represents a PTH
CC responsive gene protein of the invention.
XX
SQ Sequence 145 AA;
Query Match 100.0%; Score 778; DB 8; Length 145;
Best Local Similarity 100.0%; Pred. No. 4e-70;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MCCGSGRADAIIEPRRYESWTRTESWLTYYTDSADAPSAAPDSGPEAGLHSGMLEDEL 60
DB 1 MCCGSGRADAIIEPRRYESWTRTESWLTYYTDSADAPSAAPDSGPEAGLHSGMLEDEL 60
QY 61 PSNGVPRSTAPGCIIPNPEKKNCECTCCPNPSSISGPLETKONGIQTTEAKRDARMPAK 120
DB 61 PSNGVPRSTAPGCIIPNPEKKNCECTCCPNPSSISGPLETKONGIQTTEAKRDARMPAK 120
QY 121 EVTINVTDSIQMDRSRRITKNCVN 145
DB 121 EVTINVTDSIQMDRSRRITKNCVN 145

Db 121 EVTINVTDSIQOMDRSRRTKNCVN 145

RESULT 7
ADM46961

ID ADM46961 standard; protein; 180 AA.

XX ADM46961,

DT 03-JUN-2004 (first entry)

DE Brain and Acute Leukemia, Cytoplasmic alternate protein #3.

XX acute myelogenous leukemia; gene expression; BALC;

KM chronic myelogenous leukemia; prostate cancer; Brain and Acute Leukemia;

OS Cytoplasmic; exon.

XX Homo sapiens.

PH Key Location/Qualifiers

FT Misc-difference 41 /note= "encoded by GCS"

PN MO2003040347-A2.

PD 15-MAY-2003.

PF 12-NOV-2002; 2002WO-US036375.

PR 09-NOV-2001; 2001US-0348210P.

PA (OHIS) UNIV OHIO STATE RES FOUND.

PI Tanner SM, De la Chapel1 A;

DR WPI; 2003-441564/41.

DR N-PSDB; ADM46953.

PT Characterizing acute or chronic myelogenous leukemia, or prostate cancer

PT in a patient comprising assaying for the overexpression of one or more

PT BALC transcripts in cells obtained from the patient.

XX Disclosure; SEQ ID NO 19; 78bp; English.

XX The invention relates to a method of characterizing acute myelogenous

CC leukemia (AML) in a patient by assaying for the overexpression of one or

CC more BALC transcripts in cells obtained from the patient, where an

CC overexpression indicates that the patient has an aggressive form of AML.

CC The methods, kits and probes are useful for characterizing acute or

CC chronic myelogenous leukemia, or prostate cancer. They are also useful

CC for detecting BALC overexpression. This sequence corresponds to a BALC

CC (Brain and Acute Leukemia, Cytoplasmic) protein encoded by an alternative

CC spliced RNA consisting of exons 1, 6 and 8.

XX Sequence 180 AA:

Query Match 96.5%; Score 750.5; DB 7; Length 180;
Best Local Similarity 80.6%; Pred. No. 3.1e-67;
Matches 145; Conservative 0; Mismatches 0; Indels 35; Gaps 1;

RESULT 8
ADO48479
ID ADO48479 standard; protein; 145 AA.

XX ADO48479;

XX 12-AUG-2004 (first entry)

XX Mouse PTH responsive gene protein.

DE PTH responsive gene; PARG; bone-forming; bone; bone density modulation;

KM transgenic animal; osteopathic; gene therapy; osteoporosis; mouse;

KM murine.

OS Mus sp.

PN MO200404152-A2.

PD 27-MAY-2004.

PF 10-NOV-2003; 2003WO-US035655.

PR 12-NOV-2002; 2002US-0425532P.

PA (AMHP) WYETH.

PI Robinson JA, Stojanovic-Susulic V, Babij P, Murrills RJ;

DR WPI; 2004-420299/39.

DR N-PSDB; ADO48478.

PT New nucleic acid fragment encoding a PARG polypeptide, useful in

PT preparing a composition for diagnosing, treating or preventing bone

PT related disorders, e.g., osteoporosis.

XX Claim 9; SEQ ID NO 8; 169bp; English.

XX The invention relates to a novel PTH responsive gene (PARG) fragment

CC encoding a polypeptide. The invention further comprises: a chimeric

CC construct comprising the isolated nucleic acid fragment operatively

CC linked to suitable regulatory sequences; a host cell transformed with the

CC chimeric construct; a vector comprising the nucleic acid fragment;

CC obtaining a nucleic acid fragment encoding the polypeptide; a method for

CC fragment; an antibody that specifically binds to one or more epitopes of

CC a PARG polypeptide; a composition for regulating bone-forming activity

CC in a mammal comprising the nucleic acid fragment, polypeptide or antibody

CC ; an agent that alters the expression of PARG gene or polypeptide;

CC determining whether an agent alters the expression of PARG mRNA;

CC screening agents for effectiveness in altering expression of the nucleic

CC acid fragment; screening for agents useful for treating bone related

CC disorders; evaluating the efficacy of a treatment of a bone related

CC disorder in a subject; identifying polypeptides capable of binding to

CC PARG; monitoring the effectiveness of treatment of a subject with a bone

CC related agent; a transgenic animal comprising the DNA, an animal model

CC for the study of bone density modulation comprising a first group of

CC animals composed of the transgenic animal and a second group of control

CC animals; studying bone mass determinants; studying the modulation of bone

CC mass; studying an effect of PARG on bone disorders; identifying an agent

CC for treating bone related disorders; identifying whether an agent which

CC has bone forming activity; and a stably transfected cell line comprising

CC two constructs, the first construct comprising a ligand binding domain

CC linked to a DNA binding domain which is linked to an activation domain

CC all of which expression is driven by a constitutive promoter, the second

CC construct comprising multiple copies of DNA binding elements linked to a

CC minimal promoter which is linked to PARG cDNA, where upon the addition

CC of chemical inducer, transactivation of PARG gene is induced. The PARG

CC polynucleotide has osteopathic activity. The PTH responsive gene may be

CC used to treat disorders by gene therapy. The nucleic acid is useful in

CC preparing a composition for diagnosing, treating or preventing bone

CC related disorders, e.g., osteoporosis. This sequence represents a PTH

CC responsive gene protein of the invention.

XX Sequence 145 AA;
 SQ Query Match 83.4%; Score 649; DB 8; Length 145;
 Best Local Similarity 83.4%; Pred. No. 4e-577;
 Matches 121; Conservative 10; Mismatches 14; Indels 0; Gaps 0;
 OY 1 MCGGSRADAIERPRYESWTRETSTWLTYYTDSADAPPAAAPDSGPEAGGLHSGMLDGL 60
 DB 1 MCGGSRADAIERPRYESWTRETSTWLTYYTDSADAPPAAAPDSGPEAGGLHAGVLEDEL 60
 OY 61 PSNGVPRSTAPGGINPEKKTNCETQCPNPSGLSGPPLTKONGLOTTEAKRDKAKMPAK 120
 DB 61 SSNGVLRPAAPGGINPEKKNMCGTCPSNOSLSGPLTKONGLMTTEAKRDKAKMSAR 120
 OY 121 EVTINVTDSIQMDRSRRITKNCVN 145
 DB 121 EVAINVTENIRQMDRSKRVTKNCIN 145

RESULT 9

AD048473
 ID AD048473 standard; protein; 145 AA.

AC AD048473;
 XX

DT 12-AUG-2004 (first entry)
 XX

DE Rat PTH responsive gene protein.
 XX

KW PTH responsive gene; PAIGB; bone-forming; bone; bone density modulation;
 transgenic animal; osteopathic; gene therapy; osteoporosis; rat.
 XX

OS Rattus sp.
 XX

PN WO2004044152-A2.
 XX

PD 27-MAY-2004.
 XX

PF 10-NOV-2003; 2003WO-US035655.
 XX

PR 12-NOV-2002; 2002US-0425532P.
 XX

PA (AMHP) WYETH.
 XX

PI Robinson JA, Stojanovic-Susulic V, Babij P, Murrills RJ;
 XX

DR WPI; 2004-420299/39.
 XX

DR N-PSDB; AD048472.
 XX

PT New nucleic acid fragment encoding a PAIGB polypeptide, useful in
 preparing a composition for diagnosing, treating or preventing bone
 related disorders, e.g., osteoporosis.
 XX

XX Claim 9; SEQ ID NO 2; 169pp; English.
 PS

XX The invention relates to a novel PTH responsive gene (PAIGB) fragment
 encoding a polypeptide. The invention further comprises: a chimeric
 construct comprising the isolated nucleic acid fragment operatively
 linked to suitable regulatory sequences; a host cell transformed with the
 chimeric construct; a vector comprising the nucleic acid fragment;
 obtaining a nucleic acid fragment encoding the polypeptide; a method for
 obtaining a polypeptide; detecting the presence of the nucleic acid
 fragment; an antibody that specifically binds to one or more epitopes of
 a PAIGB polypeptide; a composition for regulating bone-forming activity
 in a mammal comprising the nucleic acid fragment, polypeptide or antibody
 ; an agent that alters the expression of PAIGB gene or polypeptide;
 CC determining whether an agent alters the expression of PAIGB mRNA;
 CC screening agents for effectiveness in altering expression of the nucleic
 acid fragment; screening for agents useful for treating bone related
 disorders; evaluating the efficacy of a treatment of a bone related
 disorder in a subject; identifying polypeptides capable of binding to
 CC PAIGB; monitoring the effectiveness of treatment of a subject with a bone

CC related agent; a transgenic animal comprising the DNA; an animal model
 CC for the study of bone density modulation comprising a first group of
 CC animals composed of the transgenic animal and a second group of control
 CC animals; studying bone mass determinants; studying the modulation of bone
 CC mass; studying an effect of PAIGB on bone disorders; identifying an agent
 CC for treating bone related disorders; identifying whether an agent which
 CC has bone forming activity; and a stably transfected cell line comprising
 CC two constructs, the first construct comprising a ligand binding domain
 CC linked to a DNA binding domain which is linked to an activation domain
 CC all of which expression is driven by a constitutive promoter, the second
 CC construct comprising multiple copies of DNA binding elements linked to a
 CC minimal promoter which is linked to PAIGB cDNA, where upon the addition
 CC of chemical inducer, transcription of PAIGB gene is induced. The PAIGB
 CC polynucleotide has osteopathic activity. The PTH responsive gene may be
 CC used to treat disorders by gene therapy. The nucleic acid is useful in
 CC preparing a composition for diagnosing, treating or preventing bone
 CC related disorders, e.g., osteoporosis. This sequence represents a PTH
 CC responsive gene protein of the invention.
 XX

XX Sequence 145 AA;
 SQ

Query Match 82.9%; Score 645; DB 8; Length 145;
 Best Local Similarity 83.4%; Pred. No. 1e-56; Mismatches 14; Indels 0; Gaps 0;
 Matches 121; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

OY 1 MCGGSRADAIERPRYESWTRETSTWLTYYTDSADAPPAAAPDSGPEAGGLHSGMLDGL 60
 DB 1 MCGGSRADAIERPRYESWTRETSTWLTYYTDSADAPPAAAPDSGPEAGGLHAGVLEDEL 60

OY 61 PSNGVPRSTAPGGINPEKKTNCETQCPNPSGLSGPPLTKONGLOTTEAKRDKAKMPAK 120
 DB 61 SSNGVLRPAAPGGINPEKKNMCGTCPSNOSLSGPLTKONGLMTTEAKRDKAKMSAR 120

OY 121 EVTINVTDSIQMDRSRRITKNCVN 145
 DB 121 EVAINVTENIRQMDRSKRVTKNCIN 145

RESULT 10
 AD046963
 ID AD046963 standard; protein; 149 AA.

AC AD046963;
 XX

DT 03-JUN-2004 (first entry)
 XX

DE Brain and Acute Leukemia, Cytoplasmic alternate protein #5.
 XX

XX acute myelogenous leukemia; gene expression; BACLC;
 KW chronic myelogenous leukemia; prostate cancer; Brain and Acute Leukemia;
 XX Cytoplasmic; exon.
 XX

OS Homo sapiens.
 XX

XX Key Location/Qualifiers
 FH Misc-difference 41
 FT /note="encoded by GCS"
 XX

PN WO2003040347-A2.
 XX

PD 15-MAY-2003.
 XX

PF 12-NOV-2002; 2002WO-US036375.
 XX

PR 09-NOV-2001; 2001US-0348210P.
 XX

PA (OHIS) UNIV OHIO STATE RES FOUND.
 XX

PI Tanner SM, De La Chapell A;
 XX

DR WPI; 2003-441564/41.
 XX

DR N-PSDB; AD046955.
 XX

PT Characterizing acute or chronic myelogenous leukemia, or prostate cancer
 PT in a patient comprises assaying for the overexpression of one or more
 PT BAALC transcripts in cells obtained from the patient.

PS Disclosure; SEQ ID NO 21; 78pp; English.

CC The invention relates to a method of characterizing acute myelogenous
 CC leukemia (AML) in a patient by assaying for the overexpression of one or
 CC more BAALC transcripts in cells obtained from the patient, where an
 CC overexpression indicates that the patient has an aggressive form of AML.
 CC The methods, kits and probes are useful for characterizing acute or
 CC chronic myelogenous leukemia, or prostate cancer. They are also useful
 CC for detecting BAALC overexpression. This sequence corresponds to a BAALC
 CC (Brain and Acute Leukemia, Cytoplasmic) protein encoded by an alternative
 CC spliced RNA consisting of exons 1, 6 and 8.

SQ Sequence 149 AA;

Query Match 73.2%; Score 569.5; DB 7; Length 149;
 Best Local Similarity 75.7%; Pred. No. 4,2e-49;
 Matches 109; Conservative 0; Mismatches 0; Indels 35; Gaps 1;

QY 1 MGGGSGRADAIERYEESWTRETESTWLTGTTSDAPPAAAPDSGPEAGGLHS----- 53
 DB 1 MGGGSGRADAIERYEESWTRETESTWLTGTTSDAPPAAAPDSGPEAGGLHSLEAKS 60
 QY 54 -----GMLDGLPSNGVPRSTARGGIPNPEKTKNCET 85
 DB 61 KIKAPDVSVDGLFASAKMAPLAVFSHQMLDGLPSNGVPRSTARGGIPNPEKTKNCET 120
 QY 86 QCPNPOSLSGGPLTORONGLOTTE 109
 DB 121 QCPNPOSLSGGPLTORONGLOTTE 144

RESULT 11

ADQ81902
 ID ADQ81902 standard; protein; 92 AA.

AC ADQ81902;

DT 09-SEP-2004 (first entry)

XX Human dihydrogenase 10.12.

XX Human; enzyme; dihydrogenase 10.12; malignant tumour; inflammation;

KW Immunological disease; haemopathy; HIV infection.

XX Homo sapiens.

XX CN1344798-A.

XX 17-APR-2002.

PF 29-SEP-2000; 2000CN-00125495.

PR 29-SEP-2000; 2000CN-00125495.

PA (SHAN-) SHANGHAI BIODOOR GENE DEV CO LTD.

PI Mao Y, Xie Y;

DR WPI; 2002-509506/55.

DR N-PSDB; ADQ81901.

PT New polypeptide human dihydrogenase 10.12 and polynucleotides encoding this
 PT polypeptide, useful for treating various diseases, such as malignant
 PT tumors, inflammations, immunological diseases, hemopathy and HIV
 PT infection.

PS Claim 1; SEQ ID NO 2; 33pp; Chinese.

XX The present invention discloses a new kind of polypeptide, human

CC dihydrogenase 10.12, polynucleotides encoding this polypeptide, a DNA
 CC recombination process to produce the polypeptide and antagonist against
 CC the polypeptide. The present invention also discloses the method of
 CC applying the polypeptide in treating various diseases, such as malignant
 CC tumors, inflammations, immunological diseases, haemopathy and HIV
 CC infection. The present sequence is the human dihydrogenase 10.12.

SQ Sequence 92 AA;

Query Match 52.1%; Score 405.5; DB 5; Length 92;
 Best Local Similarity 79.2%; Pred. No. 7.5e-33;
 Matches 80; Conservative 0; Mismatches 2; Indels 19; Gaps 1;

QY 45 GPEAGGLHSGMLEDGLPSNGVPRSTARGGIPNPEKTKNCETCCPNPOSLSGGPLTORONG 104
 DB 11 GPEAG-----NAPGIPNPEKTKNCETCCPNPOSLSGGPLTORONG 51
 QY 105 LOTTEAKRDAKMPAKEVTINTVDSIQMDRSRRITKNCVN 145
 DB 52 LOTTEAKRDAKMPAKEVTINTVDSIQMDRSRRITKNCVN 92

RESULT 12
 ADM46962
 ID ADM46962 standard; protein; 73 AA.

AC ADM46962;

DT 03-JUN-2004 (first entry)

XX Brain and Acute Leukemia, Cytoplasmic alternate protein #4.

XX acute myelogenous leukemia; gene expression; BAALC;

KW chronic myelogenous leukemia; prostate cancer; Brain and Acute Leukemia;

XX Cytoplasmic; exon.

XX Homo sapiens.

OS Key Location/Qualifiers

FT Misc-difference 41

XX MO2003040347-A2.

XX 15-MAY-2003.

PF 12-NOV-2002; 2002MO-US036375.

PR 09-NOV-2001; 2001US-0348210P.

PA (OHIS) UNIV OHIO STATE RES FOUND.

PI Tanner SM, De La Chapell A;

DR WPI; 2003-441664/41.

DR N-PSDB; ADM46954.

PT Characterizing acute or chronic myelogenous leukemia, or prostate cancer
 PT in a patient comprises assaying for the overexpression of one or more
 PT BAALC transcripts in cells obtained from the patient.

PS Disclosure; SEQ ID NO 20; 78pp; English.

CC The invention relates to a method of characterizing acute myelogenous
 CC leukemia (AML) in a patient by assaying for the overexpression of one or
 CC more BAALC transcripts in cells obtained from the patient, where an
 CC overexpression indicates that the patient has an aggressive form of AML.
 CC The methods, kits and probes are useful for characterizing acute or
 CC chronic myelogenous leukemia, or prostate cancer. They are also useful
 CC for detecting BAALC overexpression. This sequence corresponds to a BAALC
 CC (Brain and Acute Leukemia, Cytoplasmic) protein encoded by an alternative
 CC spliced RNA consisting of exons 1, 6 and 8.

SQ Sequence 73 AA;

Query Match 39.7%; Score 309; DB 7; Length 73;
 Best Local Similarity 96.6%; Pred. No. 2, 9e-23;
 Matches 56; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MCGGSGRADAIERYESWTRETSTWLTYYTSDAPPAAAPDSGPEAGGLHSGMLED 58
 |||||
 1 MCGGSGRADAIERYESWTRETSTWLTYYTSDAPPAAAPDSGPEAGGLHSGCLEE 58

RESULT 13

ADM46960
 ID ADM46960 standard; protein; 54 AA.

AC ADM46960;

DT 03-JUN-2004 (first entry)

DE Brain and Acute Leukemia, Cytoplasmic alternate protein #2.

KM acute myelogenous leukemia; gene expression; BALC;

KM chronic myelogenous leukemia; prostate cancer; Brain and Acute Leukemia;

KM Cytoplasmic; exon.

OS Homo sapiens.

FH Key Location/Qualifiers

FT Misc-difference 41 /note="encoded by GCS"

PN WO2003040347-A2.

PD 15-MAY-2003.

PF 12-NOV-2002; 2002WO-US036375.

PR 09-NOV-2001; 2001US-0348210P.

PA (OHS) UNIV OHIO STATE RES FOUND.

PI Tanner SM, De La Chapell A;

DR WPI; 2003-441564/41.

DR N-PSDB; ADM46952.

PT Characterizing acute or chronic myelogenous leukemia, or prostate cancer

PT in a patient comprises assaying for the overexpression of one or more

PT BALC transcripts in cells obtained from the patient.

PS Disclosure: SEQ ID NO 18; 78pp; English.

CC The invention relates to a method of characterizing acute myelogenous
 CC leukemia (AML) in a patient by assaying for the overexpression of one or
 CC more BALC transcripts in cells obtained from the patient, where an
 CC overexpression indicates that the patient has an aggressive form of AML.
 CC The methods, kits and probes are useful for characterizing acute or
 CC chronic myelogenous leukemia, or prostate cancer. They are also useful
 CC for detecting BALC overexpression. This sequence corresponds to a BALC
 CC (Brain and Acute Leukemia, Cytoplasmic) protein encoded by an alternative
 CC spliced RNA consisting of exons 1, 6 and 8.

SQ Sequence 54 AA;

Query Match 38.4%; Score 299; DB 7; Length 54;
 Best Local Similarity 100.0%; Pred. No. 2e-22;
 Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGGSGRADAIERYESWTRETSTWLTYYTSDAPPAAAPDSGPEAGGLHSG 54
 |||||
 1 MCGGSGRADAIERYESWTRETSTWLTYYTSDAPPAAAPDSGPEAGGLHSG 54

RESULT 14

ADO48477
 ID ADO48477 standard; protein; 54 AA.

AC ADO48477;

DT 12-AUG-2004 (first entry)

DE Human PTH responsive gene protein exon 2 splice variant.

KM PTH responsive gene; PARGB, bone-forming; bone; bone density modulation;

KM transgenic animal; osteopathic; gene therapy; osteoporosis; human.

OS Homo sapiens.

PN WO2004044152-A2.

PD 27-MAY-2004.

PF 10-NOV-2003; 2003WO-US035655.

PR 12-NOV-2002; 2002US-0425532P.

PA (AMHP) WYETH.

PI Robinson VA, Stojanovic-Susulic V, Babij P, Murrills RJ;

DR WPI; 2004-420299/39.

DR N-PSDB; ADO48476.

PT New nucleic acid fragment encoding a PARGB polypeptide, useful in

PT preparing a composition for diagnosing, treating or preventing bone

PT related disorders, e.g., osteoporosis.

PS Claim 9; SEQ ID NO 6; 16pp; English.

CC The invention relates to a novel PTH responsive gene (PARGB) fragment
 CC encoding a polypeptide. The invention further comprises: a chimeric
 CC construct comprising the isolated nucleic acid fragment operatively
 CC linked to suitable regulatory sequences; a host cell transformed with the
 CC chimeric construct; a vector comprising the nucleic acid fragment;
 CC obtaining a nucleic acid fragment encoding the polypeptide; a method for
 CC obtaining a polypeptide; detecting the presence of the nucleic acid
 CC fragment; an antibody that specifically binds to one or more epitopes of
 CC a PARGB polypeptide; a composition for regulating bone-forming activity
 CC in a mammal comprising the nucleic acid fragment, polypeptide or antibody
 CC ; an agent that alters the expression of PARGB gene or polypeptide;
 CC determining whether an agent alters the expression of PARGB mRNA;
 CC screening agents for effectiveness in altering expression of the nucleic
 CC acid fragment; screening for agents useful for treating bone related
 CC disorders; evaluating the efficacy of a treatment of a bone related
 CC disorder in a subject; identifying polypeptides capable of binding to
 CC PARGB; monitoring the effectiveness of treatment of a subject with a bone
 CC related agent; a transgenic animal comprising the DNA; an animal model
 CC for the study of bone density modulation comprising a first group of
 CC animals composed of the transgenic animal and a second group of control
 CC animals; studying bone mass determinants; studying the modulation of bone
 CC mass; studying an effect of PARGB on bone disorders; identifying an agent
 CC for treating bone related disorders; identifying whether an agent which
 CC has bone forming activity; and a stably transfected cell line comprising
 CC two constructs, the first construct comprising a ligand binding domain
 CC linked to a DNA binding domain which is linked to an activation domain
 CC all of which expression is driven by a constitutive promoter, the second
 CC construct comprising multiple copies of DNA binding elements linked to a
 CC minimal promoter which is linked to PARGB cDNA, where upon the addition
 CC of chemical inducer, transcription of PARGB gene is induced. The PARGB
 CC polynucleotide has osteopathic activity. The PTH responsive gene may be
 CC used to treat disorders by gene therapy. The nucleic acid is useful in
 CC preparing a composition for diagnosing, treating or preventing bone
 CC related disorders, e.g., osteoporosis. This sequence represents a PTH
 CC responsive gene protein of the invention.

SQ Sequence 54 AA;

Job time : 105.667 secs

Query Match 38.4%; Score 299; DB 8; Length 54;
 Best Local Similarity 100.0%; Pred. No. 2e-22;
 Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGCGSRADAIEPRYSESWTRETSTWLTYTDSAPPSAAPDGPAGGIHS 54
 |||||
 1 MCGCGSRADAIEPRYSESWTRETSTWLTYTDSAPPSAAPDGPAGGIHS 54
 Db

RESULT 15
 ADM46964
 ID ADM46964 standard; protein; 80 AA.
 XX
 AC ADM46964;
 XX

DT 03-JUN-2004 (first entry)
 XX

DE Brain and Acute Leukemia, Cytoplasmic alternate protein #6.
 XX

KM acute myelogenous leukemia; gene expression; BAALC;
 KW chronic myelogenous leukemia; prostate cancer; Brain and Acute Leukemia;
 XX Cytoplasmic; exon.
 OS Homo sapiens.
 XX

FH Key Location/Qualifiers
 FT Misc-difference 41
 FT /note= "encoded by GCS"
 XX

PN MO2003040347-A2.
 XX

PD 15-MAY-2003.
 XX

PF 12-NOV-2002; 2002WC-US036375.
 XX

PR 09-NOV-2001; 2001US-0348210P.
 XX

PA (OHIS) UNIV OHIO STATE RES FOUND.
 XX

PI Tanner SM, De la Chapell A;
 XX

DR WPI; 2003-441564/41.
 XX

DR N-PSDB; ADM46956.
 XX

PT Characterizing acute or chronic myelogenous leukemia, or prostate cancer
 PT in a patient comprises assaying for the overexpression of one or more
 PT BAALC transcripts in cells obtained from the patient.
 XX

PS Disclosure; SEQ ID NO 22; 78bp; English.
 XX

CC The invention relates to a method of characterizing acute myelogenous
 CC leukemia (AML) in a patient by assaying for the overexpression of one or
 CC more BAALC transcripts in cells obtained from the patient, where an
 CC overexpression indicates that the patient has an aggressive form of AML.
 CC The methods, kits and probes are useful for characterizing acute or
 CC chronic myelogenous leukemia, or prostate cancer. They are also useful
 CC for detecting BAALC overexpression. This sequence corresponds to a BAALC
 CC (Brain and Acute Leukemia, Cytoplasmic) protein encoded by an alternative
 CC spliced RNA consisting of exons 1, 6 and 8.
 XX

SQ Sequence 80 AA;
 XX

Query Match 37.7%; Score 293; DB 7; Length 80;
 Best Local Similarity 100.0%; Pred. No. 1.4e-21;
 Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGCGSRADAIEPRYSESWTRETSTWLTYTDSAPPSAAPDGPAGGIHS 53
 |||||
 1 MCGCGSRADAIEPRYSESWTRETSTWLTYTDSAPPSAAPDGPAGGIHS 53
 Db

Search completed: April 12, 2005, 15:37:18

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 15:15:13 ; Search time 103.667 Seconds

(without alignments)
540.967 Million cell updates/sec

Title: US-10-705-716a-2

Perfect score: 767
Sequence: 1 MCGGSSRADIEPRYESWT.....VTENIRQMRKRVTKNCIN 145

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A_GeneSeq_16Dec04:.*
1: geneseqp19808:.*
2: geneseqp19908:.*
3: geneseqp20008:.*
4: geneseqp20018:.*
5: geneseqp20028:.*
6: geneseqp20038:.*
7: geneseqp20048:.*
8: geneseqp20058:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	767	100.0	145	8 ADO48473	ADO48473 Rat PTH r
2	746	97.3	145	8 ADO48479	ADO48479 Mouse PTH
3	645	84.1	145	4 AAB95018	AAB95018 Human pro
4	645	84.1	145	5 AAO19498	AAO19498 HSI prote
5	645	84.1	145	6 ABR58646	ABR58646 Human can
6	645	84.1	145	7 ADC31800	ADC31800 Human nov
7	645	84.1	145	8 ADM46959	ADM46959 Brain and
8	645	84.1	145	7 ADO48475	ADO48475 Human PTH
9	617.5	80.5	180	7 ADM46961	ADM46961 Brain and
10	473.5	61.7	149	7 ADM46963	ADM46963 Brain and
11	335.5	43.7	92	5 ADO81902	ADO81902 Human dio
12	294	38.3	54	8 ADO48481	ADO48481 Mouse PTH
13	288	37.5	73	7 ADM46962	ADM46962 Brain and
14	278	36.2	54	7 ADM46960	ADM46960 Brain and
15	278	36.2	54	8 ADO48477	ADO48477 Human PTH
16	272	35.5	80	7 ADM46964	ADM46964 Brain and
17	140	18.3	25	7 ADM46979	ADM46979 Brain and
18	99	12.9	18	7 ADM46977	ADM46977 Brain and
19	92.5	12.1	16	8 AAR13456	AAR13456 Duffy rec
20	89	11.6	16	8 ADO48482	ADO48482 PTH respo
21	87.5	11.4	260	4 ABO9899	ABO9899 Novel hum
22	87.5	11.4	592	4 ABG15607	ABG15607 Novel hum
23	84	11.0	718	6 ABU17344	ABU17344 Protein e
24	84	11.0	726	6 ADA36828	ADA36828 Actinotro
25	83.5	10.9	700	5 AAE25052	AAE25052 Stenotrop

ALIGNMENTS

RESULT 1	ADO48473	standard; protein; 145 AA.
ID	ADO48473	
AC	ADO48473	
XX		
DT	12-AUG-2004	(first entry)
DE		Rat PTH responsive gene protein.
XX		
KM		PTH responsive gene; PAIGB; bone-forming; bone; bone density modulation;
KM		transgenic animal; osteopathic; gene therapy; osteoporosis; rat.
XX		
OS		Rattus sp.
XX		
PN	WO2004044152-A2.	
XX		
PD	27-MAY-2004.	
XX		
PF	10-NOV-2003; 2003WO-US035655.	
XX		
PR	12-NOV-2002; 2002US-0425532P.	
XX		
PA	(AMHP) WYETH.	
XX		
PI	Robinson JA, Stojanovic-Suenilic V, Babić P, Murrills RJ;	
XX		
DR	WPI; 2004-420239/39.	
XX		
PT	N-PSDB; ADO48472.	
PT	New nucleic acid fragment encoding a PAIGB polypeptide, useful in	
PT	preparing a composition for diagnosing, treating or preventing bone	
PT	related disorders, e.g., osteoporosis.	
XX		
PS	Claim 9; SEQ ID NO 2; 169pp; English.	
XX		
CC	The invention relates to a novel PTH responsive gene (PAIGB) fragment	
CC	encoding a polypeptide. The invention further comprises: a chimeric	
CC	construct comprising the isolated nucleic acid fragment operatively	
CC	linked to suitable regulatory sequences; a host cell transformed with the	
CC	chimeric construct; a vector comprising the nucleic acid fragment;	
CC	obtaining a nucleic acid fragment encoding the polypeptide; a method for	
CC	obtaining a polypeptide; detecting the presence of the nucleic acid	
CC	fragment; an antibody that specifically binds to one or more epitopes of	
CC	a PAIGB polypeptide; a composition for regulating bone-forming activity	
CC	in a mammal comprising the nucleic acid fragment, polypeptide or antibody	
CC	; an agent that alters the expression of PAIGB gene or polypeptide;	
CC	determining whether an agent alters the expression of PAIGB mRNA;	

XX New nucleic acid fragment encoding a PAlGB polypeptide, useful in
PT preparing a composition for diagnosing, treating or preventing bone
PT related disorders, e.g., osteoporosis.
XX
XX
XX Claim 9; SEQ ID NO 8; 169bp; English.
XX
XX The invention relates to a novel PTH responsive gene (PAlGB) fragment
CC encoding a polypeptide. The invention further comprises: a chimeric
CC construct comprising the isolated nucleic acid fragment operatively
CC linked to suitable regulatory sequences; a host cell transformed with the
CC chimeric construct; a vector comprising the nucleic acid fragment;
CC obtaining a nucleic acid fragment encoding the polypeptide; a method for
CC obtaining a polypeptide; detecting the presence of the nucleic acid
CC fragment; an antibody that specifically binds to one or more epitopes of
CC a PAlGB polypeptide; a composition for regulating bone-forming activity
CC in a mammal comprising the nucleic acid fragment, polypeptide or antibody
CC ; an agent that alters the expression of PAlGB gene or polypeptide;
CC determining whether an agent alters the expression of PAlGB mRNA;
CC screening agents for effectiveness in altering expression of the nucleic
CC acid fragment; screening for agents useful for treating bone related
CC disorders; evaluating the efficacy of a treatment of a bone related
CC disorder in a subject; identifying polypeptides capable of binding to
CC PAlGB; monitoring the effectiveness of treatment of a subject with a bone
CC related agent; a transgenic animal comprising the DNA; an animal model

CC animals composed of the transgenic animal and a second group of control

CC mass; studying an effect of PAIGB on bone disorders; identifying an agent

CC has bone forming activity; and a stably transfected cell line comprising

CC linked to a DNA binding domain which is linked to an activation domain

construct comprising multiple copies of DNA binding elements linked to a

CC of chemical inducer, transcription of PAIGB gene is induced. The PAIGB

used to treat disorders by gene therapy. The nucleic acid is useful in

CC related disorders, e.g., osteoporosis. This sequence represents a PTH

XX

Best Local Similarity 97.2%; Pred. No. 8.4e-74;

[illegible]

1. *Chlorophyll a* (Chl *a*)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

[illegible]

RESULT 3

ID	AAB95018 standard; protein; 145 AA.
----	-------------------------------------

AC AAB95018;

DT 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:16726.

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX	Homo sapiens.
OS	
XX	EPI074617-A2.
FN	
XX	
PD	07-FEB-2001.
XX	
PE	28-JUL-2000; 2000EP-00116126.
XX	
PR	29-JUL-1999; 99JP-00248036.
XX	
PR	27-AUG-1999; 99JP-00300253.
XX	
PR	11-JAN-2000; 2000JP-00118776.
XX	
PR	02-MAY-2000; 2000JP-00183767.
XX	
PR	09-JUN-2000; 2000JP-00241899.
XX	
PA	(HELI-) HELIX RES INST.
XX	
PI	Ota T, Isegaai T, Mishikawa T, Hayashi K, Saito K, Yamamoto J,
XX	Ishii S, Sugiyama T, Makamatsu A, Nagai K, Otsuki T;
DR	WPI; 2001-318749/34.
XX	
PT	Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT	length cDNAs defined in the specification, and for the detection and/or
PT	diagnosis of the abnormality of the proteins encoded by the full-length
PT	cDNAs.
PS	
XX	Claim 8; SEQ ID NO 16726; 2537pp + Sequence Listing; English.
CC	The present invention describes primer sets for synthesizing 5602 full-
CC	length cDNAs defined in the specification, where a primer set comprises:
CC	(a) an oligo-dT primer and an oligonucleotide complementary to the
CC	complementary strand of a polynucleotide which comprises one of the 5602
CC	nucleotide sequences defined in the specification, where the
CC	oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC	of an oligonucleotide comprising a sequence complementary to the
CC	complementary strand of a polynucleotide which comprises a 5'-end
CC	sequence and an oligonucleotide comprising a sequence complementary to a
CC	polynucleotide which comprises a 3'-end sequence, where the
CC	oligonucleotide comprises at least 15 nucleotides and the combination of
CC	the 5'-end sequence/'-end sequence is selected from those defined in the
CC	specification. The primer sets can be used in antisense therapy and in
CC	gene therapy. The primers are useful for synthesizing polynucleotides,
CC	particularly full-length cDNAs. The primers are also useful for the
CC	detection and/or diagnosis of the abnormality of the proteins encoded by
CC	the full-length cDNA. The primers allow obtaining of the full-length
CC	cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC	AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC	represent human amino acid sequences; and AAH18629 to AAH13632 represent
CC	oligonucleotides, all of which are used in the exemplification of the
CC	present invention
XX	
SQ	Sequence 145 AA;
XX	
Query Match	84.1%; Score 645; DB 4; Length 145;
Best Local Similarity	83.4%; Pred. No. 1,2e-62;
Matches 121; Conservative 10; Mismatches 14; Indels 0; Gaps 0	
DB	1 MCCGSGRADAIETPRYYESWTRETSFWLTYTSDALPSPAATDSGPGAGLAHVLEDCP 60
DB	1 MCCGSGRADAIETPRYYESWTRETSFWLTYTSDAPPSAAPPDGSGBAGLHGMLIEDGL 60
QY	61 SNGVLRPAAPGGIANPEKKMNCGTCCPNSSQSLSGPLETKOKNGLMTTEAKRPAXKMSAR 120
DB	61 PENSNGPRSTRAPGGIIPPEPKKTCTCENPPOSLSSEPLTKOKNGLOTTEAKRKDAKMPAK 120
QY	121 EVAISVTENIRQMDRSKRVTXKCIN 145
DB	121 EYTINTVDISIQQMDRERRITTKCVN 145

[illegible]

XX DE Human cancer related protein SEQ ID NO:303.
XX XX
XX KW Human; cancer; diagnosis; screening; modulator; leukemia; ischaemia;
XX KW heart disease; atherosclerosis; endometriosis.
XX OS Homo sapiens.
XX PN WO2003025138-A2.
XX PD 27-MAR-2003.
XX PF 17-SEP-2002; 2002WO-US029560.
XX PR 17-SEP-2001; 2001US-0323469P.
XX PR 20-SEP-2001; 2001US-0323887P.
XX PR 13-NOV-2001; 2001US-0350666P.
XX PR 08-FEB-2002; 2002US-0355145P.
XX PR 08-FEB-2002; 2002US-0355257P.
XX PR 12-APR-2002; 2002US-0372246P.
XX PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX PI Afar D, Aziz N, Gish KC, Hevezi PA, Mack DH, Wilson KE,
XX PI Zlotnick A;
XX DR MPI; 2003-354600/33.
XX DR N-PSDB; ACC72796.
XX PT New genes that are up-regulated or down-regulated in cancers, useful as
XX PT markers for diagnosing e.g. cancer, ischemia or heart diseases, or as
XX PT therapeutic targets for screening drugs for treating these diseases.
XX PS Claim 12; Page 753; 767pp; English.
XX CC The present invention describes an isolated nucleic acid molecule, which
XX CC comprises the sequence of any of the genes that are up-regulated or down-
XX CC regulated in specific cancers (e.g. about 1031 genes up-regulated in
XX CC acute lymphocytic leukemia). ACC72641 to ACC72860 represent cancer
XX CC related gene nucleotide sequences which encode the proteins given in
XX CC ABR8521 to ABR58709. Also described: (1) determining the presence or
XX CC absence of a pathological cell in a patient; (2) an expression vector
XX CC comprising a nucleic acid molecule described above; (3) a host cell
XX CC comprising the vector; (4) an isolated polypeptide, which is encoded by
XX CC the nucleic acid; (5) an antibody that specifically binds the polypeptide
XX CC of (4); (6) specifically targeting a compound to a pathological cell in a
XX CC patient by administering to the patient the antibody above; and (7) a
XX CC drug screening assay. The nucleic acid is useful as diagnostic markers or
XX CC therapeutic targets. In particular, the nucleic acid is useful for
XX CC diagnosing a pathology, e.g. cancer (e.g. cancer of the bone marrow,
XX CC bladder, brain, breast, cervix, colon/rectum, kidney, lung, ovary,
XX CC pancreas, prostate, skin and uterus), wounds, ischemia, heart diseases,
XX CC atherosclerosis and endometriosis. The nucleic acid is also useful in
XX CC drug screening, particularly for identifying agents for treating these
XX CC pathologies
XX SO Sequence 145 AA;
Query Match 84.1%; Score 645; DB 6; Length 145;
Best Local Similarity 83.4%; Pred. No. 1,2e-62;
Matches 121; Conservative 10; Mismatches 14; Indels 0; Gaps 0;
QY 1 MCCGSRADAIIPRYESWTRETSTWLTYYTSDALPSAATDSDGPEAGLHAGVLEDDP 60
DB 1 MCCGSRADAIIPRYESWTRETSTWLTYYTSDALPSAATDSDGPEAGLHAGVLEDDP 60
QY 61 SSGNGLRPAAPGGINPEKKNACGTCCPSOSISSLPTLTKQNGLWTTTEAKRDAXMSAR 120
DB 61 PBNNGVPRSTAPGGINPEKKNACGTCCPSOSISSLPTLTKQNGLWTTTEAKRDAXMSAR 120
QY 121 EVAISVTENIRQMSDKRTKNCIN 145
DB 121 EVTINVTDSIQMDRSRRITKNCVN 145

RESULT 6
ADCC31800
ID ADCC31800 standard, protein; 145 AA.
XX AC ADCC31800;
XX XX
XX DT 18-DEC-2003 (first entry)
XX XX
XX DE Human novel polypeptide sequence, SEQ ID NO:1882.
XX XX
XX KW Human; diagnostic; drug screening; forensics; gene mapping;
XX KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
XX KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
XX KW ulcers; osteoporosis; autoimmune disease; cancer;
XX KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
XX KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vlnetary;
XX KW antitumor; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
XX KW gene therapy; chromosome 8.
XX OS Homo sapiens.
XX PN WO2003029271-A2.
XX PD 10-APR-2003.
XX PF 24-SEP-2002; 2002WO-US030474.
XX PR 24-SEP-2001; 2001US-0324631P.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang TY, Zhang J, Ren F, Yue AJ, Zhao QA, Wang J, Wehrman T;
XX PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
XX PI Haley-Vicente D, Dmatnac RT;
XX DR MPI; 2003-371981/35.
XX DR N-PSDB; ADC30829.
XX PT New polynucleotide and polypeptide useful for diagnosing, preventing or
XX PT treating conditions such as neurodegenerative diseases, anemia, platelet
XX PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
XX PS cancer.
XX CC Claim 20; SEQ ID NO 1882; 1185bp; English.
XX CC The invention relates to 971 novel human cDNA sequences (ADC29919-
XX CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
XX CC invention also relates to nucleic acid sequences over 99% identical with
XX CC the novel human cDNAs. The invention additionally encompasses expression
XX CC vectors and host cells comprising a nucleic acid of the invention; the
XX CC recombinant production of a polypeptide of the invention; an antibody
XX CC against a polypeptide of the invention; a method of detecting
XX CC polynucleotides or polypeptides of the invention; and methods of
XX CC identifying a compound which binds to a polypeptide; an antibody
XX CC invention further discloses methods of preventing, treating or
XX CC ameliorating a medical condition; kits comprising polynucleotide probes
XX CC and/or monoclonal antibodies for carrying out the methods of the
XX CC invention; methods for the identification of compounds that modulate the
XX CC expression or activity of the polynucleotide and/or polypeptide; and 767
XX CC coding sequences corresponding to the cDNA sequences of the invention
XX CC (ADC31861-ADC32637) and the polypeptides encoded by the cDNAs (ADC32628
XX CC -ADC3394). The nucleic acids and polypeptides of the invention are
XX CC useful in diagnostics, drug screening, forensics, gene mapping, in the
XX CC identification of mutations responsible for genetic disorders or other
XX CC traits, for assessing biodiversity, and in producing many other types of
XX CC data and products dependent on DNA and amino acid sequences. They are
XX CC also used for treating diseases such as Parkinson's disease, Alzheimer's
XX CC disease and other neurodegenerative diseases, anaemia, platelet
XX CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
XX CC cancer. The nucleic acids may also be used as hybridisation probes or
XX CC primers, and in the recombinant production of a protein. The polypeptides

PT	Characterizing acute or chronic myelogenous leukemia, or prostate cancer
PT	in a patient comprises assaying for the overexpression of one or more
PT	BALC transcripts in cells obtained from the patient.
PS	Diclosure; SEQ ID NO 21; 76bp; English.
XX	
CC	The invention relates to a method of characterizing acute myelogenous
CC	leukemia (AML) in a patient by assaying for the overexpression of one or
CC	more BALC transcripts in cells obtained from the patient, where an
CC	overexpression indicates that the patient has an aggressive form of AML.
CC	The methods, kits and probes are useful for characterizing acute or
CC	chronic myelogenous leukemia, or prostate cancer. They are also useful
CC	for detecting BALC overexpression. This sequence corresponds to a BALC
CC	(Brain and Acute Leukemia, Cytoplasmic) protein encoded by an alternative
CC	spliced RNA consisting of exons 1, 6 and 8.
XX	
SEQ	Sequence 149 AA;
Query Match	61.7%; Score 473.5; DB 7; Length 149;
Best Local Similarity	66.0%; Pred. No. 1e-43;
Matches	95; Conservative 2; Mismatches 12; Indels 35; Gaps 1;
OY	1 MGCGGSRADATEPRRYESMTRETSWTLYTDSALPSAAAADSPEAGLHA----- 53
DB	1 MGCGGSRADALEPRTYESWTRTESWTLYTDSAPPAAADSPGEAGLHSLVEAEKS 60
OY	54 -----GVLEDDGPSNGVLRPAAPGGIAPNEKKMNCCT 85 :
DB	61 KIKAPTDSVDEGLFSAKMAPLAIPSHGMLEDGLPSNKRSTPAAGIIPNEKTNCET 120
OY	86 QCPNSQSLSSGPPLTKONGWLTTTE 109
DB	121 QCPNPQSLSSGPPLTKONGLCQTTE 144
RESULT 11	
ID	ADQ81902 standard; protein; 92 AA.
ADQ81902	
XX	
ADQ81902;	
XX	
DT	09-SEP-2004 (first entry)
XX	
DE	Human dioxygenase 10.12.
XX	
KW	Human; enzyme; dioxygenase 10.12; malignant tumour; inflammation;
KW	immunological disease; haemopathy; HIV infection.
XX	
OS	Homo sapiens.
XX	
CNI	CNI344798-A.
XX	
PD	17-APR-2002.
XX	
EP	29-SEP-2000; 2000CN-00125495.
XX	
PR	29-SEP-2000; 2000CN-00125495.
XX	
PA	(SHAN-) SHANGHAI BIODOOR GENE DEV CO LTD.
XX	
PI	Mao Y, Xie Y;
XX	
DR	WPI; 2002-509506/55.
XX	
N-P	N-P5DB; ADQ81901.
XX	
PT	New polypeptide human dioxygenase 10.12 and polynucleotides encoding this
PT	polypeptide, useful for treating various diseases, such as malignant
PT	tumors, inflammations, immunological diseases, hemopathy and HIV
XX	
PT	infection.
XX	
CS	Claim 1; SEQ ID NO 2; 33pp; Chinese.
XX	
XX	The present invention discloses a new kind of polypeptide, human

[illegible]

disorders; evaluating the efficacy of a treatment of a bone related disorder in a subject; identifying polypeptides capable of binding to PAIGB; monitoring the effectiveness of treatment of a subject with a bone related agent; a transgenic animal comprising the DNA; an animal model for the study of bone density modulation comprising a first group of animals composed of the transgenic animal and a second group of control animals; studying bone mass determinants; studying the modulation of bone mass; studying an effect of PAIGB on bone disorders; identifying an agent for treating bone related disorders; identifying whether an agent which has bone forming activity; and a stably transfected cell line comprising two constructs, the first construct comprising a ligand binding domain linked to a DNA binding domain which is linked to an activation domain all of which expression is driven by a constitutive promoter, the second construct comprising multiple copies of DNA binding elements linked to a minimal promoter which is linked to PAIGB cDNA, where upon the addition of chemical inducer, transcription of PAIGB gene is induced. The PAIGB polynucleotide has osteoparathic activity. The PTH responsive gene may be used to treat disorders by gene therapy. The nucleic acid is useful in preparing a composition for diagnosing, treating or preventing bone related disorders, e.g., osteoporosis. This sequence represents a PTH responsive gene protein of the invention.

SO Sequence 54 AA;

Query Match 38.3%; Score 294; DB 8; Length 54;
Best Local Similarity 100.0%; Pred. No. 1.6e-24;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGGSRADAIIPRYESWTREESTWLTYSDDALPSAAATDGPAGGLHAG 54
DB 1 MCGGSRADAIIPRYESWTREESTWLTYSDDALPSAAATDGPAGGLHAG 54

RESULT 13

ADM46962 standard; protein; 73 AA.

AC ADM46962;

DT 03-JUN-2004 (first entry)

DE Brain and Acute Leukemia, Cytoplasmic alternate protein #4.

XX acute myelogenous leukemia; gene expression; BAALC;

KM chronic myelogenous leukemia; prostate cancer; Brain and Acute Leukemia;

XX Cytoplasmic; exon.

OS Homo sapiens.

FH Key Location/Qualifiers

FT Misc-difference 41

PN MO2003040347-A2.

PD 15-MAY-2003.

PF 12-NOV-2002; 2002MO-US036375.

PR 09-NOV-2001; 2001US-0348210P.

PS (OHIS) UNIV OHIO STATE RES FOUND.

PI Tanner SM, De la Chapell A;

DR WPI, 2003-441564/41.

XX N-PSDB; ADM46954.

PT Characterizing acute or chronic myelogenous leukemia, or prostate cancer

PS in a patient comprises assaying for the overexpression of one or more

XX BAALC transcripts in cells obtained from the patient.

PS Disclosure; SEQ ID NO 20; 78pp; English.

XX The invention relates to a method of characterizing acute myelogenous leukemia (AML) in a patient by assaying for the overexpression of one or more BAALC transcripts in cells obtained from the patient, where an overexpression indicates that the patient has an aggressive form of AML. The methods, kits and probes are useful for characterizing acute or chronic myelogenous leukemia, or prostate cancer. They are also useful for detecting BAALC overexpression. This sequence corresponds to a BAALC (Brain and Acute Leukemia, Cytoplasmic) protein encoded by an alternative spliced RNA consisting of exons 1, 6 and 8.

SQ Sequence 73 AA;

Query Match 37.5%; Score 288; DB 7; Length 73;
Best Local Similarity 91.4%; Pred. No. 1.1e-23;
Matches 53; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MCGGSRADAIIPRYESWTREESTWLTYSDDALPSAAATDGPAGGLHAG 58
DB 1 MCGGSRADAIIPRYESWTREESTWLTYSDDALPSAAATDGPAGGLHAG 58

RESULT 14

ADM46960 standard; protein; 54 AA.

AC ADM46960;

DT 03-JUN-2004 (first entry)

DE Brain and Acute Leukemia, Cytoplasmic alternate protein #2.

XX acute myelogenous leukemia; gene expression; BAALC;

KM chronic myelogenous leukemia; prostate cancer; Brain and Acute Leukemia;

XX Cytoplasmic; exon.

OS Homo sapiens.

FH Key Location/Qualifiers

FT Misc-difference 41

PN MO2003040347-A2.

PD 15-MAY-2003.

PF 12-NOV-2002; 2002MO-US036375.

PR 09-NOV-2001; 2001US-0348210P.

PS (OHIS) UNIV OHIO STATE RES FOUND.

PI Tanner SM, De la Chapell A;

DR WPI, 2003-441564/41.

XX N-PSDB; ADM46952.

PT Characterizing acute or chronic myelogenous leukemia, or prostate cancer

PS in a patient comprises assaying for the overexpression of one or more

XX BAALC transcripts in cells obtained from the patient.

PS Disclosure; SEQ ID NO 18; 78pp; English.

XX The invention relates to a method of characterizing acute myelogenous leukemia (AML) in a patient by assaying for the overexpression of one or more BAALC transcripts in cells obtained from the patient, where an overexpression indicates that the patient has an aggressive form of AML.

XX The methods, kits and probes are useful for characterizing acute or chronic myelogenous leukemia, or prostate cancer. They are also useful for detecting BAALC overexpression. This sequence corresponds to a BAALC (Brain and Acute Leukemia, Cytoplasmic) protein encoded by an alternative

XX spliced RNA consisting of exons 1, 6 and 8.

SQ Sequence 54 AA;
 Query Match 36.2%; Score 278; DB 7; Length 54;
 Best Local Similarity 94.4%; Pred. No. 9.2e-23;
 Matches 51; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MCGGSRADAIEPRYYESWTRETSTWLTYYTDSDALPSAAATDSCPEAGGLHAG 54
 1 MCGGSRADAIEPRYYESWTRETSTWLTYYTDSDALPSAAATDSCPEAGGLHAG 54
 Db 1 MCGGSRADAIEPRYYESWTRETSTWLTYYTDSDALPSAAATDSCPEAGGLHAG 54
 RESULT 15
 ADO48477
 ID ADO48477 standard; protein; 54 AA.
 XX
 AC ADO48477;
 XX
 DT 12-AUG-2004 (first entry)
 XX
 DE Human PTH responsive gene protein exon 2 splice variant.
 XX
 KW PTH responsive gene; PAIGB; bone-forming; bone; bone density modulation;
 KW transgenic animal; osteopathic; gene therapy; osteoporosis; human.
 OS
 OS Homo sapiens.
 XX
 PN MO2004044152-A2.
 XX
 PD 27-MAY-2004.
 XX
 PF 10-NOV-2003; 2003MO-US035655.
 XX
 PR 12-NOV-2002; 2002US-0425532P.
 XX
 (AMBP) WYETH.
 PI Robinson JA, Stojanovic-Susulic V, Babić P, Murrills RJ;
 XX
 DR WPI; 2004-420299/39.
 DR N-PSDB; ADO48476.
 XX
 PT New nucleic acid fragment encoding a PAIGB polypeptide, useful in
 PT preparing a composition for diagnosing, treating or preventing bone
 PT related disorders, e.g., osteoporosis.
 XX
 PS Claim 9; SEQ ID NO 6; 169pp; English.
 XX
 CC The invention relates to a novel PTH responsive gene (PAIGB) fragment
 CC encoding a polypeptide. The invention further comprises: a chimeric
 CC construct comprising the isolated nucleic acid fragment operatively
 CC linked to suitable regulatory sequences; a host cell transformed with the
 CC chimeric construct; a vector comprising the nucleic acid fragment;
 CC obtaining a nucleic acid fragment encoding the polypeptide; a method for
 CC obtaining a polypeptide; detecting the presence of the nucleic acid
 CC fragment; an antibody that specifically binds to one or more epitopes of
 CC a PAIGB polypeptide; a composition for regulating bone-forming activity
 CC in a mammal comprising the nucleic acid fragment; polypeptide or antibody
 CC; an agent that alters the expression of PAIGB gene or polypeptide;
 CC determining whether an agent alters the expression of PAIGB mRNA;
 CC screening agents for effectiveness in altering expression of the nucleic
 CC acid fragment; screening for agents useful for treating bone related
 CC disorders; evaluating the efficacy of a treatment of a bone related
 CC disorder in a subject; identifying polypeptides capable of binding to
 CC PAIGB; monitoring the effectiveness of treatment of a subject with a bone
 CC related agent; a transgenic animal comprising the DNA; an animal model
 CC for the study of bone density modulation comprising a first group of
 CC animals composed of the transgenic animal and a second group of control
 CC animals; studying bone mass determinants; studying the modulation of bone
 CC mass; studying an effect of PAIGB on bone disorders; identifying an agent
 CC for treating bone related disorders; identifying whether an agent which
 CC has bone forming activity; and a stably transfected cell line comprising
 CC two constructs, the first construct comprising a ligand binding domain
 CC linked to a DNA binding domain which is linked to an activation domain

CC all of which expression is driven by a constitutive promoter, the second
 CC construct comprising multiple copies of DNA binding elements linked to a
 CC minimal promoter which is linked to PAIGB cDNA, where upon the addition
 CC of chemical inducer, transcription of PAIGB gene is induced. The PAIGB
 CC polynucleotide has osteopathic activity. The PTH responsive gene may be
 CC used to treat disorders by gene therapy. The nucleic acid is useful in
 CC preparing a composition for diagnosing, treating or preventing bone
 CC related disorders, e.g., osteoporosis. This sequence represents a PTH
 CC responsive gene protein of the invention.
 XX
 SQ Sequence 54 AA;
 Query Match 36.2%; Score 278; DB 8; Length 54;
 Best Local Similarity 94.4%; Pred. No. 9.2e-23;
 Matches 51; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MCGGSRADAIEPRYYESWTRETSTWLTYYTDSDALPSAAATDSCPEAGGLHAG 54
 1 MCGGSRADAIEPRYYESWTRETSTWLTYYTDSDALPSAAATDSCPEAGGLHAG 54
 Db 1 MCGGSRADAIEPRYYESWTRETSTWLTYYTDSDALPSAAATDSCPEAGGLHAG 54
 Search completed: April 12, 2005, 15:37:16
 Job time : 109.667 secs

This Page Blank (uspto)

```

RESULT 1
US-09-949-016-7117
; Sequence 7117, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7117
; LENGTH: 694
; TYPE: PRt
; ORGANISM: Human
US-09-949-016-7117

Query Match      12.1%; Score 92.5; DB 4; Length 694;
Best Local Similarity 31.9%; Pred. No. 0.1;
Matches 30; Conservative 7; Mismatches 32; Indels 25; Gaps 3;

QY       31 TDSALPSAATDSCPEAGGHH-----AGVLEGGPSSNGVRPAAP-- 71
          |||:|||||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db        233 TDSEVSOSPANKGSKRPHSNQHQPSPAVPPYPGSPPPASALSTTPGNNGVPAPAAPPs 292

QY       72 --GGIANPEKMCNCTCOPNSOSL---SGPLT 99
          |::|||:|||||:|||||:|||||:|||||:|||||:|||||:
Db        293 ALGPKASPAPSHNSTGPAYIAOVAVAAPPAPSGBST 326

RESULT 2
Patent No. 5198347-4
APPLICANT: MILLER, LOUIS H.; ADAMS, JOHN H.; KASLOW,
DAVID C.; FRANG, XIANGDOUG
TITLE OF INVENTION: DNA ENCODING PLASMIDIUM VIVAX AND
PLASMIDIUM KNOWLESII DUFRY RECEPTOR
NUMBER OF SEQUENCES: 27
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/554,837
FILING DATE: 2008-JUL-1990
(SEQ ID NO:4)

```

5198347-4
LENGTH: 778

Query Match 12.1%; Score 92.5; DB 6; Length 778;
Best Local Similarity 33.0%; Pred. No. 0.12;
Matches 29; Conservative 12; Mismatches 44; Indels 3; Gaps 3;

QY 31 TSDALPSAATDSCGPEAGGLHAGVLEDPSSNGV-LRPAAPGIANPEKKNK-GTQCP 88
DB 352 TVSSDVPVSGKSGSTSSHALAGENGVEHNGTTEPERDEKADPKQDIEVKKQDT 411

QY 89 NSQSLSS-GPLQKQKGLMTTEAKRDAK 115
DB 412 DRSQGLSGPHTDERATLGETHMEKDE 439

RESULT 3
5198347-4

Patent No. 5198347
APPLICANT: Miller, LOUIS H.; ADAMS, JOHN H.; KASLOW,
DAVIC C.; FANG, XIANGDOUG
TITLE OF INVENTION: DNA ENCODING PLASMIDIUM VIVAX AND
PLASMIDIUM KNOWLES DUFFY RECEPTOR
NUMBER OF SEQUENCES: 27
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/554,837
FILING DATE: 20-JUL-1990
SEQ ID NO: 4
LENGTH: 778

Query Match 12.1%; Score 92.5; DB 6; Length 778;
Best Local Similarity 33.0%; Pred. No. 0.12;
Matches 29; Conservative 12; Mismatches 44; Indels 3; Gaps 3;

QY 31 TSDALPSAATDSCGPEAGGLHAGVLEDPSSNGV-LRPAAPGIANPEKKNK-GTQCP 88
DB 352 TVSSDVPVSGKSGSTSSHALAGENGVEHNGTTEPERDEKADPKQDIEVKKQDT 411

QY 89 NSQSLSS-GPLQKQKGLMTTEAKRDAK 115
DB 412 DRSQGLSGPHTDERATLGETHMEKDE 439

RESULT 4
US-09-328-352-8115

Sequence 8115, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 8115
LENGTH: 726
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-8115

Query Match 11.0%; Score 84; DB 4; Length 726;
Best Local Similarity 35.1%; Pred. No. 1;
Matches 27; Conservative 8; Mismatches 22; Indels 20; Gaps 4;

QY 2 GCGSRADAIEPRYVSWTRETSTWLTITDSDALPSA--AATDSG-----PEAGGLHAG 54
DB 181 GFGAGGEDVWEPNDVNWGDEKE--WLAHNSALAGSNLAATTEMGLIYVNE----- 231

QY 55 VLEDGSSNGVLRPAAP 71
DB 232 ----GPOAGDPRSAAP 244

RESULT 5
US-09-408-647A-2

Sequence 2, Application US/09408647A
Patent No. 6399858
GENERAL INFORMATION:
APPLICANT: Kobayashi, Donald
TITLE OF INVENTION: Chitinase Gene from Stenotrophomonas
TITLE OF INVENTION: maltophilia
FILE REFERENCE: Rut-Cook 98-0090
CURRENT APPLICATION NUMBER: US/09/408,647A
CURRENT FILING DATE: 1999-08-26
PRIOR APPLICATION NUMBER: 60/098,036
PRIOR FILING DATE: 1998-08-27
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 700
TYPE: PRT
ORGANISM: Stenotrophomonas maltophilia

NAME/KEY: SIGNAL
LOCATION: (1)...(41)
NAME/KEY: DOMAIN
LOCATION: (196)...(290)
NAME/KEY: DOMAIN
LOCATION: (330)...(483)

PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: Genbank No. 6399858 AF014950
DATABASE ENTRY DATE: 1997-09-23
US-09-408-647A-2

Query Match 10.9%; Score 83.5; DB 3; Length 700;
Best Local Similarity 24.0%; Pred. No. 1.1;
Matches 31; Conservative 14; Mismatches 57; Indels 27; Gaps 2;

QY 26 TWLTTSDALPSAATDSCGPEAGGLHAGVLEDPSSNGVLRPAAPGIANPEKKNK-GTQCP 81
DB 159 TVANASAGSHTKAVATNNANVTSSATVSVTVTASNDTTPSPVGGIASSKATATYV 218

QY 82 -----NCG-----TCCPNSQSLSSGPLQKQKGLMTTEAKRDAKMS 118
DB 219 LVWSAATDNGSGGVAGYDVYRNGSLVSGPSATQYTDGLTASTAYTVTRARDNAGMS 278

QY 119 AREVAISVT 127
DB 279 AQSGSISVT 287

RESULT 6
US-09-071-035-176

Sequence 176, Application US/09071035
Patent No. 6448043
GENERAL INFORMATION:
APPLICANT: Gil H. Choi
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 496
CORRESPONDENCE ADDRESS:
ADDRESSER: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,035
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 176:
SEQUENCE CHARACTERISTICS:
LENGTH: 286 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-176

Query Match 10.4%; Score 79.5; DB 4; Length 286;
Best Local Similarity 26.6%; Pred. No. 0.89; Mismatches 73; Indels 41; Gaps 9;
Matches 47; Conservative 16;

QY 3 CGG-----SRADAIEPRYSEW---TRETESTWLTYSDDLPSAAAT-DSGPEAG 49
DB 1 CGGKSTENTDSSSAESTTVSTKASATKSSSK-ATTKSDAKPSGTTTADSKATAS 59
QY 50 GLHAGVLEDPSSNGVLRPAAPGIANPEKKN-CGTQCPN---SGLSSGPLYTKQNGU 105
DB 60 STKEA-----ANNGSAEKQSPAKNANPDQANOVNLQANMPGQGLPQALITSGTNMF 113
QY 106 WTEAKRDAKRMAREVAISVTENIRQMD-RSKRYT-----KNCIN 145
DB 114 LTAATTSQADQNNFRVLYAEKEALPVNDARVNQLTPISSFEKTYGSDAEAKNAVN 170

RESULT 7
US-09-071-035-174
Sequence 174, Application US/09071035
GENERAL INFORMATION:
APPLICANT: Gil H. Choi
TITLE OF INVENTION: Enterococcus faecalis polynucleotides and polypeptides
NUMBER OF SEQUENCES: 496
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,035
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 174:
SEQUENCE CHARACTERISTICS:

LENGTH: 305 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-174

Query Match 10.4%; Score 79.5; DB 4; Length 305;
Best Local Similarity 26.6%; Pred. No. 0.98; Mismatches 73; Indels 41; Gaps 9;
Matches 47; Conservative 16;

QY 3 CGG-----SRADAIEPRYSEW---TRETESTWLTYSDDLPSAAAT-DSGPEAG 49
DB 20 CGGKSTENTDSSSAESTTVSTKASATKSSSK-ATTKSDAKPSGTTTADSKATAS 78
QY 50 GLHAGVLEDPSSNGVLRPAAPGIANPEKKN-CGTQCPN---SGLSSGPLYTKQNGU 105
DB 79 STKEA-----ANNGSAEKQSPAKNANPDQANOVNLQANMPGQGLPQALITSGTNMF 132
QY 106 WTEAKRDAKRMAREVAISVTENIRQMD-RSKRYT-----KNCIN 145
DB 133 LTAATTSQADQNNFRVLYAEKEALPVNDARVNQLTPISSFEKTYGSDAEAKNAVN 189

RESULT 8
US-08-374-077C-2
Sequence 2, Application US/08374077C
Patent No. 6027912
GENERAL INFORMATION:
APPLICANT: Hall, Linda M.
APPLICANT: Ren, Dejian
APPLICANT: Zheng, Wei
APPLICANT: Dubald, Manuel Marcel Paul
TITLE OF INVENTION: Genes Encoding an Invertebrate Alpha1
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, LLP
STREET: 699 Prince Street
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22314-3187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/374,077C
FILING DATE: 19-JAN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm M.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 022650-264
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6620
TELEFAX: 703-836-2021
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2516 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-374-077C-2

Query Match 10.2%; Score 78.5; DB 3; Length 2516;
Best Local Similarity 22.4%; Pred. No. 27; Mismatches 39; Conservative 21; Indels 53; Gaps 8;
Matches 39; Conservative 21; Mismatches 39; Indels 53; Gaps 8;
QY 3 CGGSRADAIEPRY---YESWTRETESTWLTYSDDLPSAAATDSGPEAGGLHAGVLEDP 60

Db 207 CCGGGISAPPPRLTPEAMQLOPQ-----NSVTSAGSTNSPSSSGG-----GR 249
QY 61 SSGVLRPAAPGSIAPPEKKMC-----GTQC-----PNS 90
Db 250 DNNSSY--SAVGDDSSSNCSNCDITGDNSTLHGLGVGVCSFIADCDNSEDDGDPNN 307
QY 91 QSLSSGPL-TOKONGLMTTEAKRDARMSAREVAISVTENIRQMDRSKRYTNC 143
Db 308 QDLSSQTLRTAIVAAVAALAAKEQAQESLADCE-SFSDRRQDADBDVRIIQC 360

RESULT 9

US-08-895-590-2
; Sequence 2, Application US/08895590
; Patent No. 6207410
; GENERAL INFORMATION:
; APPLICANT: Hall, Linda M.
; APPLICANT: Ren, Dejian
; APPLICANT: Zheng, Wei
; APPLICANT: Dubald, Manuel Marcel Paul
; TITLE OF INVENTION: Genes Encoding an Insect Calcium Channel
; NUMBER OF SEQUENCES: 101
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, LLP
; STREET: 699 Prince Street
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22314-3187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/895,590
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/374,888
; FILING DATE: 19-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm M.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 022650-263
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6620
; TELEFAX: 703-836-2021
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2516 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-895-590-2

Query Match 10.2%; Score 78.5; DB 3; Length 2516;
Best Local Similarity 22.4%; Pred. No. 27;
Matches 39; Conservative 21; Mismatches 61; Indels 53; Gaps 8;

QY 3 CCGSRADAIEPRY--YESWTRTESTWLTYYTSDALPSAAATDSCPEAGGLHAGVLEDP 60
Db 207 CCGGGISAPPPRLTPEAMQLOPQ-----NSVTSAGSTNSPSSSGG-----GR 249
QY 61 SSGVLRPAAPGSIAPPEKKMC-----GTQC-----PNS 90
Db 250 DNNSSY--SAVGDDSSSNCSNCDITGDNSTLHGLGVGVCSFIADCDNSEDDGDPNN 307
QY 91 QSLSSGPL-TOKONGLMTTEAKRDARMSAREVAISVTENIRQMDRSKRYTNC 143
Db 308 QDLSSQTLRTAIVAAVAALAAKEQAQESLADCE-SFSDRRQDADBDVRIIQC 360

RESULT 10

US-09-539-879A-2
; Sequence 2, Application US/09539879A
; Patent No. 6436627
; GENERAL INFORMATION:
; APPLICANT: Hall, Linda M.
; APPLICANT: Ren, Dejian
; APPLICANT: Zheng, Wei
; APPLICANT: Dubald, Manuel Marcel Paul
; TITLE OF INVENTION: Genes Encoding an Invertebrate Alpha1
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, LLP
; STREET: 699 Prince Street
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22314-3187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/539,879A
; FILING DATE: 31-Mar-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/111,865
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/374,077
; FILING DATE: 19-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm M.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 022650-264
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6620
; TELEFAX: 703-836-2021
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2516 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-539-879A-2

Query Match 10.2%; Score 78.5; DB 4; Length 2516;
Best Local Similarity 22.4%; Pred. No. 27;
Matches 39; Conservative 21; Mismatches 61; Indels 53; Gaps 8;

QY 3 CCGSRADAIEPRY--YESWTRTESTWLTYYTSDALPSAAATDSCPEAGGLHAGVLEDP 60
Db 207 CCGGGISAPPPRLTPEAMQLOPQ-----NSVTSAGSTNSPSSSGG-----GR 249
QY 61 SSGVLRPAAPGSIAPPEKKMC-----GTQC-----PNS 90
Db 250 DNNSSY--SAVGDDSSSNCSNCDITGDNSTLHGLGVGVCSFIADCDNSEDDGDPNN 307
QY 91 QSLSSGPL-TOKONGLMTTEAKRDARMSAREVAISVTENIRQMDRSKRYTNC 143
Db 308 QDLSSQTLRTAIVAAVAALAAKEQAQESLADCE-SFSDRRQDADBDVRIIQC 360

RESULT 11

US-09-121-964-1
; Sequence 1, Application US/09121964
; Patent No. 6124447
; GENERAL INFORMATION:
; APPLICANT: Natori, Shunji
; TITLE OF INVENTION: NOVEL ENZYME CATALYZING DEPHOSPHORYLATION

```

; FILE REFERENCE: 32290-144753
; CURRENT APPLICATION NUMBER: US/09/121,964
; CURRENT FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 724
; TYPE: PR1
; ORGANISM: Sarcophaga peregrina
US-09-121-964-1

```

```

Query Match          10.2%; Score 78; DB 3; Length 724;
Best Local Similarity 28.6%; Pred. No. 5.1;
Matches 33; Conservative 6; Mismatches 40; Indels 34; Gaps 4;

```

```

QY 25 STWLTITDSDALPSAATDSCP-EAGGLHAGVLEDDG-----PSSNGV 65
DB 2 STTLEVTQKDLHDLLEIFKKPFDAGSCERKSLSEEDITNRCCEFLFKKDYTLIELDNGV 61
QY 66 LRPAAPGSIANPE-----KMNCGTCPCNSQSLSSGPTOKONGL 105
DB 62 LSPRYPERIRITPEYEHGHTSTTTPNNTNLGTQATNHPQSG--KNGNGL 110

```

```

RESULT 12
US-09-914-259-10
; Sequence 10, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/09/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 885
; TYPE: PR1
; ORGANISM: Mus musculus
US-09-914-259-10

```

```

Query Match          10.2%; Score 78; DB 4; Length 885;
Best Local Similarity 29.6%; Pred. No. 6.8;
Matches 34; Conservative 16; Mismatches 33; Indels 32; Gaps 7;

```

```

QY 49 GGHAGVLEDDGSPSGVLRPAAPGSI-----ANPEKKNCGTCPCNSQSLSSGP---L 98
DB 381 GSITHS---DKPT--ILRPATVGGTLEDGCTQAKKQKAPCVSE--SSAGAGPANMA 431
QY 99 TOKONGIWTTEAKR-----DAKMSAREVAISVTEN---IRQMDRSKEVT 140
DB 432 TQCKEGRYSRPSKRGPLSKLMADEGEFTSARAVLITVTKDHDHLDQFSNVAVS 486

```

```

RESULT 13
US-09-220-528-104
; Sequence 104, Application US/09220528A
; Patent No. 6284540
; GENERAL INFORMATION:
; APPLICANT: Milbrandt, Jeffrey D.
; APPLICANT: Baloh, Robert H.
; TITLE OF INVENTION: Artemin, A No. 6284540e1 Neurotrophic Factor
; FILE REFERENCE: 6029-7998
; CURRENT APPLICATION NUMBER: US/09/220,528A
; CURRENT FILING DATE: 1998-12-24
; EARLIER APPLICATION NUMBER: 09/218,698
; EARLIER FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 60/108,148
; EARLIER FILING DATE: 1998-11-12
; EARLIER APPLICATION NUMBER: 09/163,283

```

```

; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 104
; LENGTH: 215
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-220-528-104

```

```

Query Match          10.1%; Score 77.5; DB 3; Length 215;
Best Local Similarity 26.5%; Pred. No. 1;
Matches 31; Conservative 12; Mismatches 41; Indels 33; Gaps 5;

```

```

QY 5 GSRADAIER-----YYSWTRETSYTLTYTDSALPSAATDGPPEAGGHA----- 53
DB 65 GADPAALGGRGGAACARSWCRCASAWATAFTSWCVSASAPAAABA--LHTTSAMP 122
QY 54 GVLDEDPSSNGVLRPAAPGSIANPEKKNCGTCPCNSQSLSSGPTOKONGLWTTTA 110
DB 123 AYVAPPCD---RPPAPGSAAP-----AADPRATKSPSWTSTA 159

```

```

RESULT 14
US-09-620-405B-473
; Sequence 473, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuguu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C8
; CURRENT APPLICATION NUMBER: US/09/620,405B
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 495
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 473
; LENGTH: 445
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-620-405B-473

```

```

Query Match          10.1%; Score 77.5; DB 4; Length 445;
Best Local Similarity 21.5%; Pred. No. 2.9;
Matches 37; Conservative 28; Mismatches 58; Indels 49; Gaps 7;

```

```

QY 7 RADAIER-----RYESWTRETSYTLTYTDS-ALPSAATDSGPEAGGLHAGVLEDDG 60
DB 78 RADEILPSSKQDYESSWDSSELCETVSQKQVCLPKAHQKEIDKIN---GKLEESP 133
QY 61 SSGVLRPAAPGSIANPEKKN-----NCGTCPCNS 90
DB 134 DNDGFLKAPCRMKVSIPTALBLMDNQTFKAEPEKPSAFEPALIEKQSVNPALTKOE 193
QY 91 OSLSG---PLTOKONGL-----WTEAKRDAKMSAREVAISVTENIRQMDR 135
DB 194 QTLRADQMPFSESQKQKVENSGWDSLSLET--VSQKQVCPKATHTQKEMDK 243

```

```

RESULT 15
US-09-433-826B-473
; Sequence 473, Application US/09433826B
; Patent No. 6579973
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuguu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.

```

```

; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.4700C4
; CURRENT APPLICATION NUMBER: US/09/433.826B
; CURRENT FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 474
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 473
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-433-826B-473

```

Query Match	10.1%	Score 77.5	DB 4	Length 445
Best Local Similarity	21.5%	Pred. No. 2.9		
Matches 37	Conservative 28	Mismatches 58	Indels 49	Gaps 7

```

0Y 7 RAALIE-----RYSSWIRETESTLWJYDSD-ALPSAAANDSPPEAGLHAGVLEGP 60
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 78 RAELIPSESKQDYVESSMDSSELCETVSQKDVCLPKAAHQKEIDKIN---SKLRESP 133
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
0Y 61 SSGVLRPAAPGGIANPEKKM-----NCGNOCENS 90
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 134 DINGFLKAPRMKVSLPTALBIMNQTFKAPPEKPSAFEPALIMQKSVNKALIELKNE 193
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
0Y 91 QSLSSG---PLTOKONGL-----WTEPAKRDARMSAREVALSVTENIKOMR 135
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 194 QTRPADOMFPSESKQKKEVENSMDSESLRET--VSQKDCVCPKATHQKEMK 243
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

Search completed: April 12, 2005, 15:43:28
Job time : 27.6667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 15:37:30 ; Search time 75.6667 Seconds
(without alignments)
636.206 Million cell updates/sec

Title: US-10-705-716A-2

Perfect score: 767
Sequence: 1 MCGGSRADAIEPRYESWT.....VTENIRQMDRSKVTNKCIN 145

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1418010 seqs, 33197259 residues

Total number of hits satisfying chosen parameters: 1418010

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Published Applications AA:
1: /cgn2_6/prodata/1/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubppa/PCT_NEW_PUB.pep:*
3: /cgn2_6/prodata/1/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/prodata/1/pubppa/US06_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubppa/US07_NEW_PUB.pep:*
6: /cgn2_6/prodata/1/pubppa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/prodata/1/pubppa/US08_NEW_PUB.pep:*
8: /cgn2_6/prodata/1/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/prodata/1/pubppa/US09A_PUBCOMB.pep:*
10: /cgn2_6/prodata/1/pubppa/US09B_PUBCOMB.pep:*
11: /cgn2_6/prodata/1/pubppa/US09C_PUBCOMB.pep:*
12: /cgn2_6/prodata/1/pubppa/US09_NEW_PUB.pep:*
13: /cgn2_6/prodata/1/pubppa/US10_PUBCOMB.pep:*
14: /cgn2_6/prodata/1/pubppa/US10_PUBCOMB.pep:*
15: /cgn2_6/prodata/1/pubppa/US10C_PUBCOMB.pep:*
16: /cgn2_6/prodata/1/pubppa/US10_NEW_PUB.pep:*
17: /cgn2_6/prodata/1/pubppa/US10_NEW_PUB.pep:*
18: /cgn2_6/prodata/1/pubppa/US11_NEW_PUB.pep:*
19: /cgn2_6/prodata/1/pubppa/US60_NEW_PUB.pep:*
20: /cgn2_6/prodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	645	84.1	145	US-10-293-239-17	Sequence 17, Appl
2	645	84.1	145	US-10-177-390-30	Sequence 30, Appl
3	617.5	80.5	160	US-10-293-239-19	Sequence 19, Appl
4	473.5	61.7	149	US-10-293-239-21	Sequence 21, Appl
5	288	37.5	73	US-10-293-239-20	Sequence 20, Appl
6	278	36.2	54	US-10-293-239-18	Sequence 18, Appl
7	272	35.5	80	US-10-293-239-22	Sequence 22, Appl
8	140	18.3	25	US-10-293-239-37	Sequence 37, Appl
9	99	12.9	18	US-10-293-239-35	Sequence 35, Appl
10	94	12.3	147	US-10-767-701-48905	Sequence 48905, A
11	85.5	11.1	274	US-10-425-114-65588	Sequence 65588, A
12	85	11.1	450	US-10-437-963-137075	Sequence 137075, A
13	84	11.0	718	US-10-282-122A-45268	Sequence 45268, A

14	83	10.8	1433	9	US-09-801-368-60	Sequence 60, Appl
15	83	10.8	1433	15	US-10-149-310-112	Sequence 112, Appl
16	82.5	10.8	539	15	US-10-424-599-227399	Sequence 227399, A
17	82.5	10.8	609	16	US-10-437-963-144983	Sequence 144983, A
18	82	10.7	1491	16	US-10-437-963-140197	Sequence 140197, A
19	82	10.7	1597	16	US-10-437-963-180225	Sequence 180225, A
20	81.5	10.6	600	16	US-10-408-765A-1945	Sequence 1945, Ap
21	81	10.6	670	14	US-10-156-76A-14107	Sequence 14107, A
22	80.5	10.5	320	15	US-10-425-114-43318	Sequence 43318, A
23	80.5	10.5	508	16	US-10-437-963-188141	Sequence 188141, A
24	79.5	10.4	286	9	US-09-071-035-176	Sequence 176, Appl
25	79.5	10.4	286	14	US-10-206-576-176	Sequence 176, Appl
26	79.5	10.4	286	17	US-10-912-362-116	Sequence 174, Appl
27	79.5	10.4	305	9	US-09-071-035-174	Sequence 174, Appl
28	79.5	10.4	305	14	US-10-206-576-174	Sequence 174, Appl
29	79.5	10.4	305	17	US-10-912-362-114	Sequence 174, Appl
30	79.5	10.4	661	13	US-10-007-805-552	Sequence 552, Appl
31	79.5	10.4	661	14	US-10-076-622-552	Sequence 552, Appl
32	79.5	10.4	661	14	US-10-124-805-552	Sequence 552, Appl
33	79.5	10.4	1013	13	US-10-007-805-553	Sequence 553, Appl
34	79.5	10.4	1013	14	US-10-076-622-553	Sequence 553, Appl
35	79.5	10.4	1013	14	US-10-124-805-553	Sequence 553, Appl
36	78.5	10.2	262	16	US-10-424-599-193656	Sequence 193656, A
37	78	10.2	355	15	US-10-424-599-242710	Sequence 242710, A
38	78	10.2	455	15	US-10-087-192-1317	Sequence 1317, Ap
39	78	10.2	661	16	US-10-451-467A-554	Sequence 554, Appl
40	78	10.2	832	13	US-10-087-192-1317	Sequence 1317, Ap
41	78	10.2	885	14	US-10-080-608A-110	Sequence 10, Appl
42	78	10.2	885	15	US-10-370-685-99	Sequence 99, Appl
43	78	10.2	1479	16	US-10-437-963-180217	Sequence 180217, A
44	77.5	10.1	215	9	US-09-320-920-104	Sequence 104, Appl
45	77.5	10.1	445	9	US-09-825-301-29	Sequence 29, Appl

ALIGNMENTS

RESULT 1	
US-10-293-239-17	
; Sequence 17, Application US/10293239	
; Publication No. US20030119043A1	
; GENERAL INFORMATION:	
; APPLICANT: Tanner, Stephan	
; TITLE OF INVENTION: BAAC Expression as a diagnostic marker for acute leukemia	
; FILE REFERENCE: 22727/04101	
; CURRENT APPLICATION NUMBER: US/10/293,239	
; CURRENT FILING DATE: 2002-11-12	
; PRIOR APPLICATION NUMBER: US 60/346,210	
; PRIOR FILING DATE: 2001-11-09	
; NUMBER OF SEQ ID NOS: 39	
; SOFTWARE: PatentIn version 3.1	
; SEQ ID NO 17	
; LENGTH: 145	
; TYPE: PRT	
; ORGANISM: Homo sapiens	
US-10-293-239-17	
Query Match	84.1%; Score 645; DB 14; Length 145;
Best Local Similarity	83.4%; Pred. No. 7, 1e-57;
Matches 121; Conservative 10; Mismatches 14; Indels 0; Gaps 0;	
QY	1 MCGGSRADAIEPRYESWTRETSTWLTYSDDALPSAAATDSGEAGLHAGVLEDP 60
DB	1 MCGGSRADAIEPRYESWTRETSTWLTYSDDAPSAAPSCGEAGLHAGVLEDP 60
QY	61 SSGVGRAPAPGGINAPKMKMGCTCCPSQSLSSEPLTKQNGLTTEAKRPAKMSAR 120
DB	61 PSNGVGRAPAPGGINAPKMKMGCTCCPSQSLSSEPLTKQNGLTTEAKRPAKMSAR 120
QY	121 EVAISTENIRQMDRSKVTNKCIN 145
DB	121 EVTINTDSIQMDRSKVTNKCIN 145

RESULT 2

US-10-177-390-30
; Sequence 30, Application US/10177390
; Publication No. US20030143743A1
; GENERAL INFORMATION:
; APPLICANT: Schuler, Gerold
; APPLICANT: N.V. Antwerp Innovatiecentrum
; TITLE OF INVENTION: Improved Transfection of Eucaryotic Cells with linear
; FILE REFERENCE: 021505wo/JH/ml
; CURRENT APPLICATION NUMBER: US/10/177,390
; CURRENT FILING DATE: 2002-06-20
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-390-30

Query Match 84.1%; Score 645; DB 14; Length 145;
Best Local Similarity 83.4%; Pred. No. 7.1e-57;
Matches 121; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERPRYESWTRETESTWLTYSDDLPSAATDSGPEAGGLHAGVLEDP 60
DB 1 MCGGSRADAIERPRYESWTRETESTWLTYSDDLPSAATDSGPEAGGLHAGVLEDP 60
QY 61 SSNGVLRLPAPGAIANPEKKNCGTCCPNQSLSGGPLTQKONGLTTEAKRDKAKMSAR 120
DB 61 PSNGVPRSTAPGIPNPEKKNCGTCCPNQSLSGGPLTQKONGLTTEAKRDKAKMPAK 120
QY 121 EVAISTENIRQMDRSKRVTKNCIN 145
DB 121 EVTINVTDSIQMDRSRRITKNCIN 145

RESULT 3

US-10-293-239-19
; Sequence 19, Application US/10293239
; Publication No. US20030119043A1
; GENERAL INFORMATION:
; APPLICANT: Tanner, Stephan
; APPLICANT: de la Chapelle, Albert
; TITLE OF INVENTION: BALC expression as a diagnostic marker for acute leukemia
; FILE REFERENCE: 22727/04101
; CURRENT APPLICATION NUMBER: US/10/293,239
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/348,210
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-239-19

Query Match 80.5%; Score 617.5; DB 14; Length 180;
Best Local Similarity 67.2%; Pred. No. 5.5e-54;
Matches 121; Conservative 10; Mismatches 14; Indels 35; Gaps 1;

QY 1 MCGGSRADAIERPRYESWTRETESTWLTYSDDLPSAATDSGPEAGGLHAGVLEDP 53
DB 1 MCGGSRADAIERPRYESWTRETESTWLTYSDDLPSAATDSGPEAGGLHAGVLEDP 60
QY 54 -----GVLDDGSSNGVLRLPAPGAIANPEKKNCGT 85
DB 61 KTKAPDVSDEGLPSASAKAPLAVFSHGMLDGLPSNGVPRSTAPGIPNPEKKNCGT 120
QY 86 OCPNQSLSGGPLTQKONGLTTEAKRDKAKMSAREVAISVTENIRQMDRSKRVTKNCIN 145

DB 121 OCPNQSLSGGPLTQKONGLTTEAKRDKAKMPAKVITNTVDSIQMDRSRRITKNCIN 180

RESULT 4

US-10-293-239-21
; Sequence 21, Application US/10293239
; Publication No. US20030119043A1
; GENERAL INFORMATION:
; APPLICANT: Tanner, Stephan
; APPLICANT: de la Chapelle, Albert
; TITLE OF INVENTION: BALC expression as a diagnostic marker for acute leukemia
; FILE REFERENCE: 22727/04101
; CURRENT APPLICATION NUMBER: US/10/293,239
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/348,210
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-239-21

Query Match 61.7%; Score 473.5; DB 14; Length 149;
Best Local Similarity 60.0%; Pred. No. 1.4e-39;
Matches 95; Conservative 2; Mismatches 12; Indels 35; Gaps 1;

QY 1 MCGGSRADAIERPRYESWTRETESTWLTYSDDLPSAATDSGPEAGGLHAGVLEDP 53
DB 1 MCGGSRADAIERPRYESWTRETESTWLTYSDDLPSAATDSGPEAGGLHAGVLEDP 60
QY 54 -----GVLDDGSSNGVLRLPAPGAIANPEKKNCGT 85
DB 61 KTKAPDVSDEGLPSASAKAPLAVFSHGMLDGLPSNGVPRSTAPGIPNPEKKNCGT 120
QY 86 OCPNQSLSGGPLTQKONGLTTE 109
DB 121 OCPNQSLSGGPLTQKONGLTTE 144

RESULT 5

US-10-293-239-20
; Sequence 20, Application US/10293239
; Publication No. US20030119043A1
; GENERAL INFORMATION:
; APPLICANT: Tanner, Stephan
; APPLICANT: de la Chapelle, Albert
; TITLE OF INVENTION: BALC expression as a diagnostic marker for acute leukemia
; FILE REFERENCE: 22727/04101
; CURRENT APPLICATION NUMBER: US/10/293,239
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/348,210
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-239-20

Query Match 37.5%; Score 288; DB 14; Length 73;
Best Local Similarity 91.4%; Pred. No. 2.6e-21;
Matches 53; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERPRYESWTRETESTWLTYSDDLPSAATDSGPEAGGLHAGVLEDP 58
DB 1 MCGGSRADAIERPRYESWTRETESTWLTYSDDLPSAATDSGPEAGGLHAGVLEDP 58

RESULT 6

```
US-10-293-239-18
; Sequence 18, Application US/10293239
; Publication No. US20030119043A1
; GENERAL INFORMATION:
; APPLICANT: Tanner, Stephan
; APPLICANT: de la Chapelle, Albert
; TITLE OF INVENTION: BALC expression as a diagnostic marker for acute leukemia
; FILE REFERENCE: 22727/04101
; CURRENT APPLICATION NUMBER: US/10/293,239
; PRIOR FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/348,210
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-239-18

Query Match          36.2%; Score 278; DB 14; Length 54;
Best Local Similarity 94.4%; Pred. No. 1,8e-20;
Matches 51; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MCGGSGRAAIEPRYESWTRETESTWLTYYTDSALPSAAATDGPAGGLHAG 54
DB 1 MCGGSGRAAIEPRYESWTRETESTWLTYYTDSADPPSAAPDGPAGGLHSG 54

RESULT 7
US-10-293-239-22
; Sequence 22, Application US/10293239
; Publication No. US20030119043A1
; GENERAL INFORMATION:
; APPLICANT: Tanner, Stephan
; APPLICANT: de la Chapelle, Albert
; TITLE OF INVENTION: BALC expression as a diagnostic marker for acute leukemia
; FILE REFERENCE: 22727/04101
; CURRENT APPLICATION NUMBER: US/10/293,239
; PRIOR FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/348,210
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-239-22

Query Match          35.5%; Score 272; DB 14; Length 80;
Best Local Similarity 94.3%; Pred. No. 1,2e-19;
Matches 50; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MCGGSGRAAIEPRYESWTRETESTWLTYYTDSALPSAAATDGPAGGLHAG 53
DB 1 MCGGSGRAAIEPRYESWTRETESTWLTYYTDSADPPSAAPDGPAGGLHSG 53

RESULT 8
US-10-293-239-37
; Sequence 37, Application US/10293239
; Publication No. US20030119043A1
; GENERAL INFORMATION:
; APPLICANT: Tanner, Stephan
; APPLICANT: de la Chapelle, Albert
; TITLE OF INVENTION: BALC expression as a diagnostic marker for acute leukemia
; FILE REFERENCE: 22727/04101
; CURRENT APPLICATION NUMBER: US/10/293,239
; PRIOR FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/348,210
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 39
```

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 37
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-239-37

Query Match          18.3%; Score 140; DB 14; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.5e-07;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 RADAIERPRYESWTRETESTWLTYYT 31
DB 1 RADAIERPRYESWTRETESTWLTYYT 25

RESULT 9
US-10-293-239-35
; Sequence 35, Application US/10293239
; Publication No. US20030119043A1
; GENERAL INFORMATION:
; APPLICANT: Tanner, Stephan
; APPLICANT: de la Chapelle, Albert
; TITLE OF INVENTION: BALC expression as a diagnostic marker for acute leukemia
; FILE REFERENCE: 22727/04101
; CURRENT APPLICATION NUMBER: US/10/293,239
; PRIOR FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/348,210
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-239-35

Query Match          12.9%; Score 99; DB 14; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.0049;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 DAIEPRYESWTRETEST 26
DB 1 DAIEPRYESWTRETEST 18

RESULT 10
US-10-767-701-48905
; Sequence 48905, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21 (53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; PRIOR FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 48905
; LENGTH: 147
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3476-020-P1-X1-A12.pep
US-10-767-701-48905

Query Match          12.3%; Score 94; DB 16; Length 147;
Best Local Similarity 27.5%; Pred. No. 0.22;
Matches 30; Conservative 17; Mismatches 52; Indels 10; Gaps 3;

QY 36 LPSAAATDSGPAGGLHAGVLEDEGSSNGVLRPAAPGIANREKKNKNCOTCPNSQS----
```

Db 29 LPAALPEDAGEAGADNGCAHGRPHRG-----QABAGAGARGALHCPGPHSDSCIG 84
QY 93 --LSSGPELQKQ-NGLWTEAKRDARMAAREVAISVTENIRQMDSKR 138
Db 85 RMTSTGTSGHRPCSGILSVQLRQMPRSSAEQVREKIKQIOQRREVR 133

RESULT 11

US-10-425-114-65588
; Sequence 65588, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 65588
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4863-011-Fl12_Fli1.pep
US-10-425-114-65588

Query Match 11.1%; Score 85.5; DB 15; Length 274;
Best Local Similarity 32.4%; Pred. No. 3.4;
Matches 35; Conservative 10; Mismatches 42; Indels 21; Gaps 5;

QY 32 DSDALPSAATDSCPGAGGILAGVLEDDGSSNGVLRPAPAGI---ANPEKKNCGTQCP 88
Db 147 DSDDAAPAAAAHDDQPAVAAMAGLGGSSSG-LPPAAGAATAAEPSLSLSGLPLP 205

QY 89 -----NSQSSSGPELQKQNGLWTEAKRDARMAAREV 122
Db 206 AAEPAEAADDESRRNQGAAS-PLIEEGEG--NAQLAVRRVREEV 250

RESULT 12

US-10-437-963-137075
; Sequence 137075, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 137075
; LENGTH: 450
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_38593C.1.pep
US-10-437-963-137075

Query Match 11.1%; Score 85; DB 16; Length 450;
Best Local Similarity 25.6%; Pred. No. 7.1;
Matches 45; Conservative 14; Mismatches 53; Indels 64; Gaps 8;

QY 4 GGSRADAI-----EPRYESWTRETESTLTYTSDALPSAATDSCPGAGGIL 52
Db 3 GGS-ADAVTKEMEALLVQGNPNVAVSGETCETSSKKGKVAADSGSHSSPPEDDDAOG-- 59

QY 53 AGVLEDGSSNGVLRPAPAGGIANPEKKNCGTQCPNSQSLSGPELQ----- 100
Db 60 -----DGSQDG-----GSEPAKKKKK-----SKSKKKGPLQOTDPPSITIDELF 101

QY 101 -----KONGLM--TTEAKRDARMAAREVAISVTENIRQMDRS 136
Db 102 PSGDPPEBGIQGYKODNLMRTTSEKRELERLQKPMYNAVRAAEVHRQVRHRS 157

RESULT 13

US-10-282-122A-45268
; Sequence 45268, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 45268
; LENGTH: 718
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-10-282-122A-45268

Query Match 11.0%; Score 84; DB 15; Length 718;
Best Local Similarity 35.1%; Pred. No. 16;
Matches 27; Conservative 8; Mismatches 22; Indels 20; Gaps 4;

QY 2 GCGSGRADAIAPRYESWTRETESTLTYTSDALPSA--AATDSC-----PEAGGLHAG 54

Db 173 GFGAGEDVWEPDNDVWNGDEKE--WLAHNSBALAGSNIAATEMGLIYVNP----- 223
QY 55 VLEDGSSNGVLRPAAP 71
Db 224 ----GFGAGSDPRSAP 236

RESULT 14
US-09-801-368-60
; Sequence 60, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Bueby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Maxon, Kevin
; APPLICANT: Milne, Todd
; APPLICANT: No. US20020128250A1man, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 60
; LENGTH: 1433
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-60

Query Match 10.8%; Score 83; DB 9; Length 1433;
Best Local Similarity 29.4%; Pred. No. 49;
Matches 32; Conservative 15; Mismatches 50; Indels 12; Gaps 4;
QY 36 LPSAATDSGPAGGLHAGVLEDGSSNGVLRPAAPGIANPEKXNCGTQCPSQSLS 95
Db 927 LPATTSLKPLFGSQSKSLNRQRTPNPKR-----ENPEHEYLGNDSNNNNSEA 979
QY 96 G--PLTQKONGLMTTEAKRDKMSAREVAISTENIR--QMDRSKRV 140
Db 980 GHSFMTNTTNGKRLKYEKDKR-NAKDGISKGENAHNFQNDTKKMS 1027

RESULT 15
US-10-149-310-112
; Sequence 112, Application US/10149310
; Publication No. US20040077039A1
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas
; APPLICANT: Madden, Kevin T.
; APPLICANT: Maxon, Mary
; APPLICANT: Sherman, Amir
; TITLE OF INVENTION: Modulation of Secondary Metabolite Production by
; FILE REFERENCE: 14184-019US1
; CURRENT APPLICATION NUMBER: US/10/149,310
; CURRENT FILING DATE: 2003-02-19
; PRIOR APPLICATION NUMBER: PCT/US01/29288
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: US 60/233,564
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 308

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 112
; LENGTH: 1433
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-149-310-112

Query Match 10.8%; Score 83; DB 15; Length 1433;
Best Local Similarity 29.4%; Pred. No. 49;
Matches 32; Conservative 15; Mismatches 50; Indels 12; Gaps 4;
QY 36 LPSAATDSGPAGGLHAGVLEDGSSNGVLRPAAPGIANPEKXNCGTQCPSQSLS 95
Db 927 LPATTSLKPLFGSQSKSLNRQRTPNPKR-----ENPEHEYLGNDSNNNNSEA 979
QY 96 G--PLTQKONGLMTTEAKRDKMSAREVAISTENIR--QMDRSKRV 140
Db 980 GHSFMTNTTNGKRLKYEKDKR-NAKDGISKGENAHNFQNDTKKMS 1027

Search completed: April 12, 2005, 16:03:42
Job time : 79.6667 secs

this Page Blank (uspto)

Result No.	Score	Query Match	Length	DB	ID	Description
1	92.5	12.1	571	2	T43456	hypothetical prote
2	92.5	12.1	778	2	A35970	erythrocyte-bindin
3	88	11.5	365	1	GNVSSC	genome polypeptide
4	83	10.8	1433	2	S54587	CAR8 protein - yeast
5	81	10.6	269	2	T37073	hypothetical protei
6	81	10.6	967	2	S66952	hypothetical protei
7	79.5	10.4	2559	2	T093144	hypothetical protei
8	78.5	10.2	1139	2	A40670	probable guanine n
9	78	10.2	885	2	T09225	nuclear envelope f
10	78	10.2	3488	2	T34418	A kinase anchor pr
11	76.5	10.0	600	2	C69899	hypothetical prote
12	76	9.9	645	2	S19156	conserved hypothec
13	76	9.9	839	2	B84874	serotonin receptor
14	75.5	9.8	499	2	S22571	hypothetical prote
15	75.5	9.8	1063	2	T03743	integrase-like prote
16	75.5	9.8	1122	2	T47424	bifocal protein -
17	75	9.8	368	1	TVMSTM	hypothetical protei
18	74.5	9.8	521	2	S15163	transforming protei
19	74.5	9.7	601	2	AH0784	XpOlycomb - Africa
20	74.5	9.7	4957	2	T03455	probable transpor
21	74.5	9.7	5262	2	T03454	ALR protein - huma
22	74	9.6	260	2	B3854	ALR protein - huma
23	74	9.6	742	2	T38001	tropoin 1 - fruit
24	74	9.6	832	2	T31878	probable phosphati
25	74	9.6	962	2	T00262	hypothetical protei
26	73.5	9.6	876	1	A57988	hypothetical protei
27	73.5	9.6	960	2	T37916	regulatory protein
28	73.5	9.6	1575	2	S68448	probable heteroch
29	73	9.5	403	2	H98327	enantiomer-selecti

30	73	9.5	453	2	AE2955	glutathemyl-cRNA amid
31	73	9.5	781	2	T41551	hypochromic proteol
32	73	9.5	1274	2	T37193	enamelin matrix pro
33	72.5	9.5	251	2	C75521	cytochrome-related
34	72.5	9.5	284	2	I51172	transcription factor
35	72.5	9.5	435	2	AG1028	preplin [imported
36	72.5	9.5	503	2	T35053	probable epoxide-hi
37	72.5	9.5	516	2	JB0301	linulnae (EC 3.2.
38	72.5	9.5	1343	2	AF0611	cell division prot
39	72.5	9.5	3623	2	T06618	intrinsic factor-B
40	72.5	9.5	5327	2	T13564	microtubule-associ
41	72	9.4	108	2	T31565	hypothetical prote
42	72	9.4	1038	1	UC5757	DNA-directed DNA p
43	72	9.4	1038	2	T18222	DNA polymerase del
44	72	9.4	1573	2	T50113	3-dehydroquinase s
45	71.5	9.3	381	2	AB3048	8-amino-7-oxononan

ALIGNMENTS

RESULT 1
 T43456
 Hypothetical protein DKFZp434L061.1 - human
 C,Species: Homo sapiens (man)
 C,Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2000
 A,Accession: T43456
 R,Pouetka, A.; Kleib, M.; Mewes, H.W.; Gaassenhuber, J.; Wiemann, S.
 Submitted to the Protein Sequence Database, December 1999
 A,Reference number: Z22516
 A,Accession: T43456
 A,Status: Preliminary
 A,Molecule type: mRNA
 A,Residues: 1-571 >AA>
 A,Cross-references: UNIPROT:O75175; EMBL:AL133647
 A,Experimental source: adult testis; clone DKFZp434L061
 C,Genetics:
 A,Note: DKFZp434L061.1
 A,Superfamily: hydroxyproline-rich glycoprotein

Query Match	12.1%	Score 92.5;	DB 2;	length 571;
Best Local Similarity	31.9%	Pred. No. 0.95;		
Matches 30; Conservative	7;	Mismatches 32;	Indels 25;	Gaps 3

```

QY      3 TQSDALPSAAATDSCGEAGGLH-----AGVLEGPSSNVLPAP--  71
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      111 TQSEVSGSPAKNGSKRVHSNQHPOSRAVPPTPTSPGPPPAASALSTTPGNNVPAAPAPS  176
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY      72 --GGINPEKKNCGTQCPNSQSLS----SGPLT  99
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      171 ALGPKKASPAASHNSGTPATVAQVAPAPSGEST  204
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

RESULT 2
A35970
erythrocyte-binding protein - Plasmodium knowlesi
C/Species: Plasmodium knowlesi
C/Date: 14-Dec-1990 #sequence_revision 14-Dec-1990 #ext_change 09-Jul-2004
C/Accession: A35970
R/Adams, J.H.; Hudson, D.E.; Torti, M.; Ward, G.E.; Wellems, T.E.; Atkawa, M.; Miller, I.
Cell 63, 141-153, 1990
A/Title: The Duffy receptor family of plasmodium knowlesi is located within the micronem
A/Reference number: A35970; MUID:91004213; PMID:2170017
A/Accession: A35970
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-778 <ADA>
A/Cross-references: UNIPROT:P22545; GB:M68518; GB:M37513; NID:g160273; PID:g160274
;keywords: transmembrane protein

Query Match	12.1%;	Score 92.5;	DB 2;	Length 778;
Best Local Similarity	33.0%;	Pred. No. 1.4;		
Matches	29;	Conservative	12;	Mismatches 44;
				Indels 3;
				Gaps 3

R:Gallion, L.; Dujon, B.
 Submitted to the Protein Sequence Database, July 1996
 A/Reference number: S66854
 A/Accession: S66854
 A/Molecule type: DNA
 A/Residues: 1-967 <GA1>
 A/Cross-references: EMBL:Z74897; NID:G1420063; PID:e21930; PID:G1420064; MIPS:YOL155c
 A/Experimental source: strain S288C
 R:Gamo, F.J.; Lafuente, M.J.; Casamayor, A.; Aldea, M.; Casae, C.; Arto, J.; Herrero, E.
 Submitted to the EMBL Data Library, July 1995
 A/Description: Analysis of the DNA sequence of a 15500 bp fragment of the left arm of the
 pen reading frames.
 A/Reference number: S67324
 A/Accession: S67325
 A/Molecule type: DNA
 A/Residues: 1-164, 'GTSTTSSGSSSSTSSSVSGSTGTSATSSGSSASGS', 166-186, 'V', 188-967 <GAM>
 A/Cross-references: EMBL:X69715; NID:G1177620; PID:e190152; PID:G1177622
 R:Gamo, F.J.; Lafuente, M.J.; Casamayor, A.; Arino, J.; Aldea, M.; Casae, C.; Herrero, E.
 Yeast 12, 709-714, 1996
 A/Title: Analysis of the DNA sequence of a 15,500 bp fragment near the left telomere of
 and two new open reading frames.
 A/Reference number: S70380; MUID:96405919; PMID:8810044
 A/Accession: S70380
 A/Status: nucleic acid sequence not shown
 A/Molecule type: DNA
 A/Residues: 55-164, 'GTSTTSSGSSSSTSSSVSGSTGTSATSSGSSASGS', 166-186, 'V', 188-286, 745-773, 781
 A/Cross-references: EMBL:X69715
 C/Genetic: G
 A/Cross-references: SGD:S0005515
 A/Map position: 15L
 A/Note: YOL155c

Query Match 10.6%; Score 81; DB 2; Length 967;
 Best Local Similarity 24.8%; Pred. No. 20;
 Matches 29; Conservative 23; Mismatches 55; Indels 10; Gaps 2;
 17 ESWTRETSTWLTYSDDALPSAAATDSCPEAGS---LHAGVLEDPSSNGVLRPAAPG 72
 Db 182 ESSGSSAGSSSSTSSSVSGSSSVSGSSSSTSSSVSGSSSVSGSSSVSGSSSVSGSS 241
 QY 73 GIANPEKMNCGTQCPNSQSLSSGPLTKQKNGLWTEAKDKARMSAREVALSVTEN 129
 Db 242 SSASESITQSGTASGSSASSTGSGSVTQSSSVSGSSAS-----SAPGSSSITPOS 292

RESULT 7
 T09144
 Probable guanine nucleotide exchange factor RhogEF2 - fruit fly (*Drosophila melanogaster*)
 N/Alternate names: Shap pel/DRHOGEF2
 C/Species: *Drosophila melanogaster*
 C/Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
 C/Accession: T09144; T09223
 R/Hecker, U.; Perrimon, N.
 Submitted to the EMBL Data Library, October 1997
 A/Reference number: Z16586
 A/Accession: T09144
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-2559 <HAE>
 A/Cross-references: UNIPROT:O44113; EMBL:AF031330; NID:G2687355; PID:G2687356
 R:Barrett, K.; Lepplin, M.; Seftelman, J.
 Cell 91, 905-915, 1997
 A/Title: The Rho GTPase and a putative the rhogEF mediate a signaling pathway for the cell
 A/Reference number: Z16618; MUID:98088790; PMID:9428514
 A/Accession: T09223
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-676, 'U', 678-837, 'U', 839-889, 'S', 891-1243, 'D', 1245-1358, 'E', 1360-1368, 'R', 1
 A/Cross-references: EMBL:AF032870; NID:G2760367; PID:G2760368
 C/Genetic: G
 A/Note: rhogEF2
 A/Cross-references: FlyBase:FBgn0023172
 A/Map position: 2; 53F1-2

A/Note: orchestrates cell shape changes during gastrulation
 C/Function: mediates actin rearrangements required for cell shape changes during gastrulation
 A/Description: rat Munc13-3 protein; protein kinase C zinc-binding repeat homology
 C/Superfamily: rat Munc13-3 protein; protein kinase C zinc-binding repeat homology
 C/Keywords: embryo; GTP exchange; signal transduction
 F/1151-1200/Domain: protein kinase C zinc-binding repeat homology <KZN>

Query Match 10.4%; Score 79.5; DB 2; Length 2559;
 Best Local Similarity 28.3%; Pred. No. 84;
 Matches 30; Conservative 11; Mismatches 56; Indels 9; Gaps 3;
 5 GSRADAEPRRYESWTRETSTWLTYSDDALPSAAATDSCPEAGGLHAGVLEDPSSNG 64
 Db 799 GSSPDNMHRHRDRIRKTTSGSW-EVEKDGSSPPGTPPPYLLSSHTVLEDPENNR 857
 QY 65 VLRPAAPG-GIANPEKMNCGTQCPNSQSLSSGPL-----TQK 102
 Db 858 GAAAGPGVFIESHQTPMAAGASSPPIPLSHNMMAQSNDTQKE 903

RESULT 8
 A40670
 Nuclear envelope protein POM 121 - rat
 C/Species: *Rattus norvegicus* (Norway rat)
 C/Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004
 C/Accession: A40670
 R/Halberg, E.; Wozniak, R.W.; Blobel, G.
 J. Cell Biol. 122, 513-521, 1993
 A/Title: An integral membrane protein of the pore membrane domain of the nuclear envelope
 A/Reference number: A40670; MUID:93328754; PMID:8335683
 A/Accession: A40670
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-1199 <HAL>
 A/Cross-references: UNIPROT:P52591; GB:Z21513; NID:G396746; PID:CAV9725.1; PID:G39674.1;
 F:803-807, 845-849, 956-960, 1010-1014, 1047-1051, 1076-1080/Region: pentapeptide motif (X-P-

Query Match 10.2%; Score 78.5; DB 2; Length 1199;
 Best Local Similarity 25.2%; Pred. No. 43;
 Matches 39; Conservative 19; Mismatches 56; Indels 41; Gaps 6;
 20 TRETESWLTYSDDALPSA---AATDSCPEAGGLHAGVLEDPSSNGVLRPAAPGIA- 75
 Db 363 SRTSSVSLSTICTGRTIPSSRNATISYSTRGVQLMKRSGPTSPSSPSSSSQTP 422
 QY 76 -NPEKMNCGTQCPNSQSLSSGPL-----TQKNGLWTT-----EAK 111
 Db 423 EEPKATKTR--EEEPCHGSSSAPLVTDKESPEKVTDPATGKQQLMTSPPTPGSSGQRK 480
 QY 112 RDAKRMASR-----EVALSVTENIRQMDR 135
 Db 481 RKIQLPSSRRGQQLTPPPPELGYSITADLDMER 515

RESULT 9
 T09225
 A kinase anchor protein AKAP-KL isoform 1 - mouse
 C/Species: *Mus musculus* (house mouse)
 C/Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
 C/Accession: T09225; T09226; T09227
 R/Bong, F.; Feldmesser, M.; Casadevall, A.; Rubin, C.S.
 J. Biol. Chem. 273, 6533-6541, 1998
 A/Title: Molecular characterization of a cDNA that encodes six isoforms of a novel murine
 A/Reference number: Z16620; MUID:98165844; PMID:9497889
 A/Accession: T09225
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-885 <DON>
 A/Cross-references: UNIPROT:O54931; EMBL:AF033274; NID:G2852696; PID:G2852697
 A/Note: binds the regulatory subunits (RII) of protein kinase AII isoforms
 A/Accession: T09226
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA

A:Residues: 1-848,862-885 <D02>
 A:Cross-references: EMBL:AF033275; NID:g2852698; PID:g2852699
 A:Note: binds the regulatory subunit (RI) of protein kinase AII isoforms
 A:Accession: T09227
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-789, 'PQGHTG' <D03>
 A:Cross-references: EMBL:AF033276; NID:g2852700; PID:g2852701
 A:Note: binds the regulatory subunit (RI) of protein kinase AII isoforms
 C:Genetics:
 A:Gene: AKAP-XL
 C:Keywords: alternative splicing; kidney; lung; signal transduction

Query Match 10.2%; Score 78; DB 2; Length 885;
 Best Local Similarity 29.6%; Pred. No. 34;
 Matches 34; Conservative 16; Mismatches 33; Indels 32; Gaps 7;

QY 49 GGIHAGVLEDPSSNGVLRPAAGT-----ANPEKKNCGTQCPNSQSLSSGP---L 98
 DB 381 GSIHS-----DKPPT--ILRPATVGTLEDGTOAAKEQKAPCVSE--SQSAGAGPANA 431

QY 99 TOKONGMTTEAKR-----DAKMSAREVAISVTEN---IRQMSRKRYT 140
 DB 432 TQKEGPPISPPSKRGPLSKLMADGERTSARAVLTWVKDHDGILDPSSRSVNV 486

RESULT 10
 T34418
 hypothetical protein F12F3.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T34418
 R:Fulton, B.; Wohldmann, P.
 Submitted to the EMBL Data Library, July 1998
 A:Description: The sequence of C. elegans cosmid F12F3.
 A:Reference number: Z21521
 A:Accession: T34418
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-3488 <F01>
 A:Cross-references: EMBL:U00022; PID:AAC25885.1; GSPDB:GN00023; CESP:F12F3.3
 A:Experimental source: strain Bristol N2; clone F12F3
 C:Genetics:
 A:Gene: CESP:F12F3.3
 A:Map position: 5
 A:Introns: 281/3; 332/1; 562/3; 600/3; 1866/3; 1944/3; 3393/1

Query Match 10.2%; Score 78; DB 2; Length 3488;
 Best Local Similarity 24.3%; Pred. No. 1.6e+02;
 Matches 33; Conservative 23; Mismatches 36; Indels 44; Gaps 7;

QY 24 ESTWLT--TYTDSALPSAATDSGP-----EAGGLHAGVLEDPSSNGVLRPAAGCIAN 76
 DB 2772 DSDMLEIANTDRNKKFKDRSLTSGEYVYQVTATGHA-----VSSPESEETN 2817

QY 77 PEKKKNCGTQCPNSQSLSSGPTLQKONGMTTEAKRDARMSAREVAI-----S 125
 DB 2818 PYKILVPGSEMPASK-----TEKK-----TDPAKSESGKAEIVAKQYDQGSAS 2866

QY 126 VTENIRQMSRKRYT 141
 DB 2867 TTEAVEE-KTKTKVKV 2881

RESULT 11
 C69899
 conserved hypothetical protein yobL - Bacillus subtilis
 C:Species: Bacillus subtilis
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
 C:Accession: C69899
 R:Kumet, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter, C.; Bron, S.; Brouillet, S.; Bruch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choc, A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Erington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Filtz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallazzi, A.; Gallier, J.; Harwood, C.R.; Hentut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Huhlo, M.P.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kunita, K.; Lapidus, A.; Lardinois, A.; Muthers, Leuber, U.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Maeda, S.; Mausel, Y., M.; Ogawa, K.; Ogawa, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon, A.; Authors: Schlicht, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror, Akuch, M.; Tamakoshi, A.; Tanaka, T.; Terpestra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K.
 A:Authors: Yoshikawa, H.F.; Zumestein, E.; Yoshikawa, H.; Danchin, A.
 A>Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
 A:Reference number: A69580; MUID:198044033; PMID:9384377
 A:Accession: C69899
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-600 <KUN>
 A:Cross-references: UNIPROT:O34330; GB:Z99114; GB:AL009126; NID:g2634230; PID:CA13792.
 C:Genetics:
 A:Experimental source: strain 168
 A:Gene: yobL

Query Match 10.0%; Score 76.5; DB 2; Length 600;
 Best Local Similarity 23.1%; Pred. No. 30;
 Matches 39; Conservative 23; Mismatches 56; Indels 51; Gaps 7;

QY 9 DAIEPRYESSWTRE-----TESTWLTGYTDSALPSAATDSGPEAGLHAGVLEDPSS 61
 DB 342 DAISAIRESYQKDVNGDAYSRSRYT-----ALGSVAVAVVGRK-----AGALINKADA 393

QY 62 SNGVLRPAAGGIA-----NPEKKNCGTQCP---NSQSLSSGPT----- 99
 DB 394 AGKVINKASQAKKIKDVIPDLVPYKPKYKALADVNPYVVDGONLKNELTNAKKIP 453

QY 100 -----QKONGMTTEAKRDARMSAREVAISVTENIRQMSRKRY 139
 DB 454 DGRKRFQTKKSPPLNKEKYD-----AYEIGKVKAKKQKVDVSRV 497

RESULT 12
 S19156
 serotonin receptor 2B - fruit fly (Drosophila melanogaster)
 N:Alternate names: 5-hydroxytryptamine receptor 2B (5-HT2B)
 C:Species: Drosophila melanogaster
 C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 09-Jul-2004
 C:Accession: S19156; S18154
 R:Saudou, F.; Bouchert, U.; Amlaiky, N.; Plasat, J.L.; Hen, R.
 EMBO J. 11, 7-17, 1992
 A>Title: A family of Drosophila serotonin receptors with distinct intracellular signals
 A:Reference number: S19156; MUID:92155185; PMID:1310937
 A:Accession: S19156
 A:Molecule type: mRNA
 A:Residues: 1-645 <SAT>
 A:Cross-references: UNIPROT:P28286; EMBL:Z11490; NID:g7506; PID:CAA77571.1; PID:g7507
 C:Genetics:
 A:Gene: FLYBase:5-HT1B
 A:Cross-references: FLYBase:FBgn0004572
 A:Superfamily: octopamine receptor type I
 C:Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
 F:124-145/Domain: transmembrane #status predicted <TM1>
 F:156-177/Domain: transmembrane #status predicted <TM2>
 F:193-214/Domain: transmembrane #status predicted <TM3>
 F:234-256/Domain: transmembrane #status predicted <TM4>
 F:284-305/Domain: transmembrane #status predicted <TM5>
 F:564-587/Domain: transmembrane #status predicted <TM6>
 F:597-619/Domain: transmembrane #status predicted <TM6>

Query Match 9.9%; Score 76; DB 2; Length 645;
 Best Local Similarity 31.0%; Pred. No. 36;
 Matches 31; Conservative 12; Mismatches 31; Indels 26; Gaps 4;

QY 29 TYTDSALPSAATDSGPEAGLHAGVLEDPSSNGVLRPAAG-----IANPEKKNCG 84

Db 489 TTPPEKALSGAGTVAGAVAGSGSGSGEAGCTEGKNAGVGLGVLASTIANPHOKL--- 545
 QY 85 TOCPNSQSLSSGPLTOKNGLMTTEAKRKDAKMSAREVAI 124
 Db 546 -----AKRRL-LEAKRE--RKAQTLAI 566

RESULT 13

E84824
 hypothetical protein At2g40040 [imported] - Arabidopsis thaliana
 C:/Species: Arabidopsis thaliana (mouse-ear cress)
 C:/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
 C:/Accession: E84824
 R:/Lin. X.; Kaul, S.; Rounsaey, S.D.; Shea, T.P.; Bentto, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Varshen, S.E.; Umayam, L.; Tallon, L.;
 euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A:/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:/Reference number: AB4420; MUID:20083487; PMID:10617197
 A:/Accession: E84824

A:/Status: preliminary
 A:/Molecule type: DNA
 A:/Residues: 1-839 <STO>
 A:/Cross-references: UNIPROT:O04207; GB:AB002093; NID:g2088657; PIDN:AAB95289.1; GSPDB:GN
 C:/Genetics:
 A:/Gene: At2g40040
 A:/Map position: 2

Query Match 9.4%; Score 76; DB 2; Length 839;
 Best Local Similarity 23.8%; Pred. No. 49;
 Matches 40; Conservative 14; Mismatches 62; Indels 52; Gaps 6;

QY 5 GSRADAIERYRYESWTR---ETESTWLTITDSDALPSAAA----- 41
 Db 338 GSGAGVGLG-----WKKSETESNGATWGSXGAAWNSWDKNIETDSEPAWG 392
 QY 42 -----TDSGEAGGL---HAGVLEDGPSSNGV-----LRPAAGGIANPEKKM 81
 Db 393 SCKKNSKSTESGPAAGWAMPKKSETEFGAGWGMGDKNSSETELGPAMGMDKKSDT 452
 QY 82 NCQTQCPNSQSLSSGPLTOKNGLMTTEAKRKDAKMSAREVAISVTEN 129
 Db 453 KSGPAAMGSTDAAAMGSSDKNN---SETSDAAAMGSRNKKTSIES 496

RESULT 14

S22571
 integrase-like protein PE65 - rat
 C:/Species: Rattus norvegicus (Norway rat)
 C:/Date: 31-Dec-1993 #sequence_revision 02-Aug-1994 #text_change 09-Jul-2004
 C:/Accession: S22571; S22572
 R:/Dullio, A.; Zambreno, N.; Mogavero, A.R.; Ammendola, R.; Cimino, F.; Russo, T.
 Nucleic Acids Res. 19, 5269-5274, 1991
 A:/Title: A rat brain mRNA encoding a transcriptional activator homologous to the DNA dir
 A:/Reference number: S22571; MUID:92020215; PMID:1323810

A:/Accession: S22571
 A:/Molecule type: mRNA
 A:/Residues: 1-499 <DU11>
 A:/Cross-references: UNIPROT:Q99MK3; EMBL:X60469; NID:g57559; PIDN:CAA42999.1; PID:g57560
 A:/Accession: S22572
 A:/Status: nucleic acid sequence not shown; translation not shown
 A:/Molecule type: DNA
 A:/Residues: 31-318 <DU12>
 A:/Cross-references: EMBL:X60468; NID:g57561; PIDN:CAA42998.1; PID:g1177617
 A:/Note: this sequence was submitted to the EMBL Data Library, July 1991
 C:/Genetics:
 A:/Intons: 88/3; 107/3; 135/3; 157/3; 207/3; 250/2; 252/2; 290/3
 C:/Keywords: transcription regulation
 P:/42-78/Domain: WW repeat homology <WW1>

Query Match 9.8%; Score 75.5; DB 2; Length 499;
 Best Local Similarity 25.6%; Pred. No. 30;
 Matches 34; Conservative 17; Mismatches 59; Indels 23; Gaps 5;

QY 13 PRYVESWTESTWLTITDSDALPSAAATDSGEAGGLHAGVLEDGPSSNGVLRPAAG 72
 Db 77 PQGNSPOESOLTTTGF-----AHQGFEEGER-----WDEPEEAPMEL----- 118
 QY 73 GIAPPEKKKNCGTQCPNSQSLSSGPLTOKNGLMTTEAKRKDAKMSAREVA-ISVTENIR 131
 Db 119 GLKDPPE-----CTLFPSSQSLSPFVPOEENLPGRNNPGLKCPAVNSLGVNMTBEEL 174
 QY 132 QMDRSKRVTKNCI 144
 Db 175 AGRSSVAVANNCI 187

RESULT 15

T03743
 bifocal protein - fruit fly (Drosophila melanogaster)
 C:/Species: Drosophila melanogaster
 C:/Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
 C:/Accession: T03743
 R:/Bahr, S.M.; Yang, X.Y.; Chia, W.
 Mol. Cell. Biol. 17, 5521-5529, 1997
 A:/Title: The Drosophila bifocal gene encodes a novel protein which colocalises with acti
 A:/Reference number: Z15048; MUID:97415628; PMID:9271427
 A:/Accession: T03743

A:/Status: preliminary; translated from GB/EMBL/DBJ
 A:/Molecule type: mRNA
 A:/Residues: 1-1063 <BAH>
 A:/Cross-references: UNIPROT:O16125; EMBL:AF011793; NID:g2388667; PIDN:AAB69991.1; PID:g2
 C:/Genetics:
 A:/Cross-references: FlyBase:FBgn0014133
 A:/Note: bifocal

Query Match 9.8%; Score 75.5; DB 2; Length 1063;
 Best Local Similarity 20.7%; Pred. No. 71;
 Matches 39; Conservative 23; Mismatches 59; Indels 67; Gaps 6;

QY 5 GSRADAIERYRYESWTR---ETESTWLTITDSDAL-----PSAAAT- 42
 Db 83 GAIDPTEPATISTTSQKRNMI GSEBSEKSIISNTSDSTGCHHSVAVASLPDAAATT 142
 QY 43 -----DSGEAGGLHAGVLEDGPSSNGVLRPAAG 72
 Db 143 NNTVPIPKQSSSLNTRSGEREMRYILSESGERDGLSEGEQPAAGVSNRCGEVETG 202
 QY 73 GIAPPEKKKNCGTQCPNSQSLSSGPLTOKNGLMTTEAKRKDAKMSAREVAISVTENIRQ 132
 Db 203 TIGSPSSSAN--QNPNPWHLK---TKCKPGQSVAEGRPSAKE-----TIVNRSKS 247
 QY 133 MDRSKRV 140
 Db 248 CSKTKSIS 255

Search completed: April 12, 2005, 15:44:43
 Job time : 26.6667 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 15:16:23, Search time 94.6667 Seconds

(Without alignments)
784.347 Million cell updates/sec

Title: US-10-705-716a-2

Sequence: 1 MCGGSRADAIERPRYRESWT.....VTENIRQMDRSKRVTKNCIN 145

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Uniprot 03: *
1: uniprot_sprot: *
2: uniprot_tramb1: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	767	100.0	145	Q920K5	Q920K5 rattus norv
2	746	97.3	145	Q8VH1	Q8VH1 mus musculu
3	645	84.1	145	Q9HA3	Q9HA3 homo sapien
4	634	82.7	145	Q8WNS9	Q8WNS9 sus scrofa
5	617.5	80.5	180	Q8WXS3	Q8WXS3 homo sapien
6	473.5	61.7	149	Q8WXS1	Q8WXS1 homo sapien
7	459	59.8	123	Q9CY89	Q9CY89 mus musculu
8	294	38.3	54	Q8VBS8	Q8VBS8 mus musculu
9	294	38.3	54	Q790N3	Q790N3 rattus norv
10	288	37.5	73	Q8WXS0	Q8WXS0 homo sapien
11	278	36.2	54	Q8WTP6	Q8WTP6 homo sapien
12	272	35.5	80	Q8WXS2	Q8WXS2 homo sapien
13	267	34.8	54	Q8WNS8	Q8WNS8 sus scrofa
14	161	21.0	32	Q801V5	Q801V5 brachydanto
15	92.5	12.1	613	Q6ZM6	Q6ZM6 homo sapien
16	92.5	12.1	753	CNO3_HUMAN	CNO3_HUMAN
17	92.5	12.1	1073	PVDA_PLAKN	PVDA_PLAKN
18	88	11.5	365	POLG_SUNVS	POLG_SUNVS
19	87	11.3	1035	Q76C74	Q76C74 saccharomyc
20	87	11.3	1713	Q8TGE1	Q8TGE1 saccharomyc
21	86	11.2	585	Q7U229	Q7U229 rhodospirill
22	86	11.2	1070	PVDG_PLAKN	PVDG_PLAKN
23	84.5	11.0	534	Q96SA2	Q96SA2 homo sapien
24	84	11.0	155	Q6USF5	Q6USF5 plasmodium
25	83.5	10.9	214	Q8MTC2	Q8MTC2 leucophaea
26	83.5	10.9	677	Q6A015	Q6A015 mus musculu
27	83.5	10.9	700	Q30678	Q30678 xanthomonas
28	83.5	10.9	751	CNO3_MOUSE	CNO3_MOUSE
29	83	10.8	561	Q89EL1	Q89EL1 bradyrhizob
30	83	10.8	581	P89204	P89204 sugarcane m
31	83	10.8	1433	CAT8_YEAST	P93113 saccharomyc

32	82.5	10.8	514	2	Q7SE2	Q7SE2 neurospora
33	82.5	10.8	853	1	AOP2_MOUSE	AOP2_MOUSE mus musculu
34	82.5	10.8	1597	2	Q6KXFO	Q6KXFO debrayomyce
35	82	10.7	339	2	Q8KRB5	Q8KRB5 streptomyce
36	82	10.7	882	1	AREA_ASPNG	AREA_ASPNG
37	82	10.7	1449	2	Q7XPB1	Q7XPB1 aspergillus
38	81.5	10.6	397	2	Q9AKA7	Q9AKA7 phaseolus v
39	81.5	10.6	852	1	AOP2_HUMAN	AOP2_HUMAN
40	81.5	10.6	888	2	Q8W0U4	Q8W0U4 borghum bic
41	81.5	10.6	3064	2	Q8B388	Q8B388 sugarcane m
42	81.5	10.6	3105	2	Q8U219	Q8U219 cercopithec
43	81	10.6	228	2	Q91TQ3	Q91TQ3 tupaiid her
44	81	10.6	269	2	Q9S1Y3	Q9S1Y3 streptomyce
45	81	10.6	670	2	Q828T2	Q828T2 streptomyce

ALIGNMENTS

RESULT 1	Q920K5	PRELIMINARY;	PRT;	145 AA.
ID	Q920K5			
AC	Q920K5;			
DT	01-DEC-2001 (TrEMBLrel. 19, Created)			
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)			
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)			
DE	Dam-A20-4 (BAALC isoform 1-6-8).			
GN	Name=dem-A20-4; Synonyms=BaalC;			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RA	Wang X., Tian Q., Li W., Okano A., Suzuki T.;			
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.			
RJ	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Sprague-Dawley;			
RC	MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;			
RA	Tanner S.M., Austin J.L., Leone G., Rush J.E., Plasse C., Heinonen K.,			
RA	Knucila S., Koltz J.E., Archer K.J.,			
RA	Califuri M.A., Bloomfield C.D., de la Chapelle A.;			
RT	"BAALC, the human member of a novel mammalian neuroectoderm gene			
RT	lineage, is implicated in hematopoiesis and acute leukemia.";			
RT	Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).			
DR	EMBL; AB073318; BAB70507.1; -.			
DR	EMBL; AF371321; AAL50517.1; -.			
DR	Pfam; PF06989; BAALC_N; 1.			
SQ	SEQUENCE 145 AA; 15475 MW; DSA27AD67456F341 CRC64;			
Query Match	100.0%; Score 767; DB 2; Length 145;			
Best Local Similarity	100.0%; Pred. No. 2.8e-61;			
Matches	145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 MCGGSRADAIERPRYRESWTRETSWLTYSDDALPSAATSGSPAGGLHGVLEDP 60			
DB	1 MCGGSRADAIERPRYRESWTRETSWLTYSDDALPSAATSGSPAGGLHGVLEDP 60			
QY	61 SSNGVLRPAAPGGIANPEKMGCTCPNSQSLSGPTLTKONGLTTEAKRDKMSAR 120			
DB	61 SSNGVLRPAAPGGIANPEKMGCTCPNSQSLSGPTLTKONGLTTEAKRDKMSAR 120			
QY	121 EVAISVTENIRQMDRSKRVTKNCIN 145			
DB	121 EVAISVTENIRQMDRSKRVTKNCIN 145			
RESULT 2	Q8VH1	PRELIMINARY;	PRT;	145 AA.
ID	Q8VH1			
AC	Q8VH1;			
DT	01-MAR-2002 (TrEMBLrel. 20, Created)			

DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DE 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DB BAIAC isoform 1-6-8 (Mus musculus 16 days neonate cerebellum cDNA,
 DE RIKEN full-length enriched library, clone:963002H16 product:brain and
 DN acute leukemia, cytoplasmic, full insert sequence).
 GN Name=Baalc;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
 RX MEDLINE=21574583; PubMed=11707601; DOI=10.1073/pnas.241525498;
 RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Pias C., Heinonen K.,
 RA Mrozek K., Sill H., Knutti S., Kolitz J.E., Archer K.J.,
 RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.,
 RT BAIAC, the human member of a novel mammalian neuroectoderm gene
 RT lineage, is implicated in hematopoiesis and acute leukemia.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.,
 RT "High-efficiency full-length cDNA cloning.";
 RL Meth. Enzymol. 303:19-44(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
 RA The FANTOM Consortium.
 RT the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=2049374; PubMed=11042159; DOI=10.1101/gr.145100;
 RX Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.,
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:11617-1630(2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
 RX MEDLINE=20530913; PubMed=11078661; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Kono H., Akiyama J., Niehi K., Kitsuai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsunaga S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:11757-1771(2000).
 RN [7]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto C., Hiraoaka T., Hirozane T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,

RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ono M., Ohsaka N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai K., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF371320; AL50516.1; -;
 DR EMBL; AK079337; BAC37611.1; -;
 DR MGD; MG1:1928704; Baalc.
 DR Pfam; PF06989; Baalc_N; 1.
 SQ SEQUENCE 145 AA; 15515 MW; 4972670A618C4D6D CRC64;
 QY Query Match 97.3%; Score 746; DB 2; Length 145;
 Best local Similarity 97.2%; Pred. No. 2,2e-59;
 Matches 141; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 DB 1 MCGGSRADAIIPRYESWTRETSWTLYTDSALPSAAATDSGPAGLHAGVLEDP 60
 1 MCGGSRADAIIPRYESWTRETSWTLYTDSALPSAAATDSGPAGLHAGVLEDP 60
 QY 61 SENGVLPPAPGIANPEKKNCGTQCPNSGSLSSGPTLQKNGIMTTEAKRDARMSAR 120
 DB 61 SENGVLPPAPGIANPEKKNCGTQCPNSGSLSSGPTLQKNGIMTTEAKRDARMSAR 120
 QY 121 EVAISYENIRQMDRSKRYTKNCIN 145
 DB 121 EVAISYENIRQMDRSKRYTKNCIN 145
 RESULT 3
 O9HA93 PRELIMINARY; PRT; 145 AA.
 AC O9HA93;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Hypothetical protein FL12015 (BAALC isoform 1-6-8) (Brain and acute
 DE leukemia, cytoplasmic) (BAALC 1-6-8).
 GN Name=BAALC;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Whole embryo;
 RX PubMed=14702039; DOI=10.1038/ng1285;
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
 RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 RA Yamamoto J., Saiki K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,
 RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
 RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
 RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
 RA Niinomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
 RA Tanai H., Kimachi M., Watanabe M., Hirooka S., Chiba Y., Ishida S.,
 RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
 RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Imamura Y.,
 RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
 RA Musashiro K., Yunkai H., Oshima A., Sasaki N., Aotsuka S.,
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara N., Sano S.,
 RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiro M., Omori Y.,
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Ohtani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
 RA Matsunuma K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
 RA Togsashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,

RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Maeno Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isegai T., Sugano S.,
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";
RL Nat. Genet. 36:40-45(2004).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plass C., Heimonen K.,
RA Mrozek K., Sill H., Knutti S., Kolitz J.E., Archer K.J.,
RA Calligaris M.A., Bloomfield C.D., de la Chapelle A.;
RT "BALC, the human member of a novel mammalian neuroectoderm gene
RT lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.U., Uebin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Ritchard S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywicki M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,
RT Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RU EMBL; AK020207; BAB13960.1; -
DR EMBL; AF311319; AAL50515.1; -
DR EMBL; BC011517; AAL11517.1; -
DR EMBL; AF363578; AAL50377.1; -
DR InterPro; IPR009728; BALC_N.
DR Pfam; PF06989; BALC_N; 1.
SQ SEQUENCE 145 AA; 15551 MW; CFB92BBE83DD92E CRC64;

Query Match 84.1%; Score 645; DB 2; Length 145;
Best Local Similarity 83.4%; Pred. No. 2,7e-50;
Matches 121; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

QY 1 MCGGSRADAIIPRYESWTRETSTWLTYSADALPSAATDSGPEAGGLHAGVLEDP 60
DB 1 MCGGSRADAIIPRYESWTRETSTWLTYSADALPSAATDSGPEAGGLHAGVLEDP 60
QY 61 SSGNGVRPAPAGINPEKKNCGTCPCNSQSISSGSLPQKNGLTTEKRAKAKMSAR 120
DB 61 SSGNGVRPAPAGINPEKKNCGTCPCNSQSISSGSLPQKNGLTTEKRAKAKMSAR 120
QY 121 EVAIVTENIRQMDRSKRVTKNCIN 145
DB 121 EVAIVTENIRQMDRSKRVTKNCIN 145

RESULT 4
Q8WNE9 PRELIMINARY; PRT; 145 AA.
AC Q8WNE9;

DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE BALC isoform 1-6-8.
GN Name=BALC;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plass C., Heimonen K.,
RA Mrozek K., Sill H., Knutti S., Kolitz J.E., Archer K.J.,
RA Calligaris M.A., Bloomfield C.D., de la Chapelle A.;
RT "BALC, the human member of a novel mammalian neuroectoderm gene
RT lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AF371322; AAL50518.1; -
DR Pfam; PF06989; BALC_N; 1.
SQ SEQUENCE 145 AA; 15401 MW; C75ED7D00EF82E26 CRC64;

Query Match 82.7%; Score 634; DB 2; Length 145;
Best Local Similarity 82.8%; Pred. No. 2,6e-49;
Matches 120; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

QY 1 MCGGSRADAIIPRYESWTRETSTWLTYSADALPSAATDSGPEAGGLHAGVLEDP 60
DB 1 MCGGSRADAIIPRYESWTRETSTWLTYSADALPSAATDSGPEAGGLHAGVLEDP 60
QY 61 SSGNGVRPAPAGINPEKKNCGTCPCNSQSISSGSLPQKNGLTTEKRAKAKMSAR 120
DB 61 SSGNGVRPAPAGINPEKKNCGTCPCNSQSISSGSLPQKNGLTTEKRAKAKMSAR 120
QY 121 EVAIVTENIRQMDRSKRVTKNCIN 145
DB 121 EVAIVTENIRQMDRSKRVTKNCIN 145

RESULT 5
Q8WXS3 PRELIMINARY; PRT; 180 AA.
ID Q8WXS3;
AC Q8WXS3;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE BALC 1-5-6-8.
GN Name=BALC;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plass C., Heimonen K.,
RA Mrozek K., Sill H., Knutti S., Kolitz J.E., Archer K.J.,
RA Calligaris M.A., Bloomfield C.D., de la Chapelle A.;
RT "BALC, the human member of a novel mammalian neuroectoderm gene
RT lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AF363578; AAL50379.1; -
DR Pfam; PF06989; BALC_N; 1.
SQ SEQUENCE 180 AA; 19224 MW; 380183E0F188F84 CRC64;

Query Match 80.5%; Score 617.5; DB 2; Length 180;
Best Local Similarity 67.2%; Pred. No. 1e-47;
Matches 121; Conservative 10; Mismatches 14; Indels 35; Gaps 1;

QY 1 MCGGSRADAIIPRYESWTRETSTWLTYSADALPSAATDSGPEAGGLHAGVLEDP 53
DB 1 MCGGSRADAIIPRYESWTRETSTWLTYSADALPSAATDSGPEAGGLHAGVLEDP 53


```
QY 54 -----GVLEDPSSNGVLRPAAPGGIANPEKXNCGT 85
DB 61 KIKAPDVSDEGLFSAKMAPLAVFSGHLEDPSSNGVLRPAAPGGIANPEKXNCGT 120
QY 86 QCPNSSLSSGPTOKONGMTTEAKRDARMSARVAISVTENINQMDRSKRYTKNCIN 145
DB 121 QCPNSSLSSGPTOKONGMTTEAKRDARMSARVAISVTENINQMDRSKRYTKNCIN 180

RESULT 6
Q8KX51 PRELIMINARY; PRT; 149 AA.
ID Q8KX51
AC Q8KX51
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DE 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DB BAALC 1-5-6-7-8.
GN Name=BAALC;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
(1)
RP SEQUENCE FROM N.A.
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plasc C., Heimonen K.,
RA Mrozek K., Sill H., Knuutila S., Kolitz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BAALC, the human member of a novel mammalian neuroectoderm gene
RT lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AF363578; AAL50381.1; -.
DR Genev; HGNC:14333; BAALC;
PFam; PF06989; BAALC.N; 1.
SQ SEQUENCE 149 AA; 15577 MW; DB26C1969B91F6E CRC64;

Query Match 61.7%; Score 473.5; DB 2; Length 149;
Best Local Similarity 66.0%; Pred. No. 7.5e-35;
Matches 95; Conservative 2; Mismatches 12; Indels 35; Gaps 1;

QY 1 MCGGSGRAIPRYVESWTRETESTWLTYYTSDALPSAANDSGEAGLHA----- 53
DB 1 MCGGSGRAIPRYVESWTRETESTWLTYYTSDAPPSAAPSGEAGLHVAEAKS 60
QY 54 -----GVLEDPSSNGVLRPAAPGGIANPEKXNCGT 85
DB 61 KIKAPDVSDEGLFSAKMAPLAVFSGHLEDPSSNGVLRPAAPGGIANPEKXNCGT 120
QY 86 QCPNSSLSSGPTOKONGMTTE 109
DB 121 QCPNSSLSSGPTOKONGMTTE 144

RESULT 7
Q9CYS9 PRELIMINARY; PRT; 123 AA.
ID Q9CYS9
AC Q9CYS9
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length
DE enriched library, clone:2810457D07 product:brain and acute leukemia,
DE cytoplasmic, full insert sequence.
GN Name=Baalc;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
(1)
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=whole body;
RC MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;

RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN (2)
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=whole body;
RC MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN (3)
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=whole body;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN (4)
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=whole body;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN (5)
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=whole body;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Alzawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama Y., Nishi K., Kitsuaki T., Tashiro H., Itoh M.,
RA Sumi N., Iehi Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Wachihi K.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN (6)
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=whole body;
RA Adachi J., Alzawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukuishi Y., Furuno M.,
RA Hangaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Kono H., Kouda M., Koya S., Kurihara C.,
RA Matsumura Y., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK013358; BAB28808.1; -.
DR MGD; MGI:1928704; Baalc.
SQ SEQUENCE 123 AA; 13391 MW; 33DFCB5F8CFAA0AB CRC64;

Query Match 59.8%; Score 459; DB 2; Length 123;
Best Local Similarity 94.6%; Pred. No. 1.2e-33;
Matches 88; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 53 AGVLEDPSSNGVLRPAAPGGIANPEKXNCGTQCPNSSLSSGPTOKONGMTTEAKR 112
DB 31 SGVLEDPSSNGVLRPAAPGGIANPEKXNCGTQCPNSSLSSGPTOKONGMTTEAKR 90
QY 113 DAKMSARVAISVTENINQMDRSKRYTKNCIN 145
DB 91 DAKMSARVAISVTENINQMDRSKRYTKNCIN 123

RESULT 8
```


Q8VB88
ID Q8VB88 PRELIMINARY; PRT; 54 AA.
AC Q8VB88;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE BAAIC isoform 1-8.
GN Name=BAAIC;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=12956/SvEvTac;
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plase C., Heinenen K.,
RA Mrozek K., Sill H., Knuutila S., Kolitz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BAAIC, the human member of a novel mammalian neuroectoderm gene
RT lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AF371324; AAL50520.1; -.
DR MGD; MGI:1928704; Baaic.
DR Pfam; PF06989; BAAIC N; 1.
SQ SEQUENCE 54 AA; 5667 MW; FE88C0BFB0DB829 CRC64;

Query Match 38.3%; Score 294; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 3.3e-19;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGGSRADAIEPRYESTWRETSTWLTYSDDALPSAATDGPAGGLHAG 54
DB 1 MCGGSRADAIEPRYESTWRETSTWLTYSDDALPSAATDGPAGGLHAG 54

RESULT 9
0790N3
ID 0790N3 PRELIMINARY; PRT; 54 AA.
AC 0790N3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE BAAIC isoform 1-8.
GN Name=BAAIC;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plase C., Heinenen K.,
RA Mrozek K., Sill H., Knuutila S., Kolitz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BAAIC, the human member of a novel mammalian neuroectoderm gene
RT lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AF371325; AAL50521.1; -.
DR Interpro; IPR009728; BAAIC_N.
DR Pfam; PF06989; BAAIC N; 1.
SQ SEQUENCE 54 AA; 5667 MW; FE88C0BFB0DB829 CRC64;

Query Match 38.3%; Score 294; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 3.3e-19;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGGSRADAIEPRYESTWRETSTWLTYSDDALPSAATDGPAGGLHAG 54
DB 1 MCGGSRADAIEPRYESTWRETSTWLTYSDDALPSAATDGPAGGLHAG 54

RESULT 10
Q8WXS0
ID Q8WXS0 PRELIMINARY; PRT; 73 AA.
AC Q8WXS0;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE BAAIC 1-4-5-6-8.
GN Name=BAAIC;
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plase C., Heinenen K.,
RA Mrozek K., Sill H., Knuutila S., Kolitz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BAAIC, the human member of a novel mammalian neuroectoderm gene
RT lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AF363578; AAL50380.1; -.
DR Pfam; PF06989; BAAIC N; 1.
SQ SEQUENCE 73 AA; 7871 MW; 98DBC2E6E6F524A CRC64;

Query Match 37.5%; Score 288; DB 2; Length 73;
Best Local Similarity 91.4%; Pred. No. 1.6e-18;
Matches 53; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MCGGSRADAIEPRYESTWRETSTWLTYSDDALPSAATDGPAGGLHAG 58
DB 1 MCGGSRADAIEPRYESTWRETSTWLTYSDDALPSAATDGPAGGLHAG 58

RESULT 11
Q8WTP6
ID Q8WTP6 PRELIMINARY; PRT; 54 AA.
AC Q8WTP6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE BAAIC isoform 1-8 (BAAIC protein) (BAAIC 1-8).
GN Name=BAAIC;
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plase C., Heinenen K.,
RA Mrozek K., Sill H., Knuutila S., Kolitz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BAAIC, the human member of a novel mammalian neuroectoderm gene
RT lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RX TISSUE=Brain;
RA Strausberg R.L., Feingold E.A., Grove L.H., Derge J.G., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,
RA Shapirostein M., Soares T.B., Bonaldo A.F., Casavant T.L., Schetz T.E.,
RA Rana S.S., Loughell N.A., Peters G.J., Abramson R.D., Millar C.S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultik S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
FAhey J., Helton E., Ketterman M., Madan A., Rodigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalak U., Smallos D.E., Schenck A., Schein J.E.,
 RA Jones S.J., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Strausberg R.,
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF371323; AAL50519.1; -;
 DR EMBL; BC035038; AAL50308.1; -;
 DR EMBL; AF363578; AAL50378.1; -;
 DR Pfam; PF06989; BAALC N; 1.
 SQ SEQUENCE 54 AA; 5663 MW; FEF0B4EABED9B825 CRC64;

Query Match 36.2%; Score 278; DB 2; Length 54;
 Best Local Similarity 94.4%; Pred. No. 9e-18;
 Matches 51; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MCGGSRADAIIPRYESWTRETESTWLTYYTDSALPSAAATDSCPEAGGLHAG 54
 DB 1 MCGGSRADAIIPRYESWTRETESTWLTYYTDSADAPPSAAAPDSCPEAGGLHAG 54

RESULT 12
 ID 08WXS2 PRELIMINARY; PRT; 80 AA.
 AC 08WXS2;
 DT 01-MAR-2002 (TRENBLrel. 20, Created)
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
 DE BAALC 1-2.
 GN Name=BAALC;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OK NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
 RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plase C., Heimonen K.,
 RA Mrozek K., Sill H., Knuutila S., Koltz J.E., Archer K.J.,
 RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.,
 RT "BAALC, the human member of a novel mammalian neuroectoderm gene
 RT lineage, is implicated in hematopoiesis and acute leukemia.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
 DR EMBL; AF363578; AAL50382.1; -;
 DR Pfam; PF06989; BAALC N; 1.
 SQ SEQUENCE 80 AA; 8573 MW; 70BE9B91C3245E2D CRC64;

Query Match 35.5%; Score 272; DB 2; Length 80;
 Best Local Similarity 94.3%; Pred. No. 5e-17;
 Matches 50; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MCGGSRADAIIPRYESWTRETESTWLTYYTDSALPSAAATDSCPEAGGLHAG 53
 DB 1 MCGGSRADAIIPRYESWTRETESTWLTYYTDSADAPPSAAAPDSCPEAGGLHAG 53

RESULT 13
 ID 08WNE8 PRELIMINARY; PRT; 54 AA.
 AC 08WNE8;
 DT 01-MAR-2002 (TRENBLrel. 20, Created)
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
 DE BAALC isoform 1-8.
 GN Name=BAALC;
 OS Sus scrofa (Pig).
 RN [1]

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OK NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
 RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plase C., Heimonen K.,
 RA Mrozek K., Sill H., Knuutila S., Koltz J.E., Archer K.J.,
 RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.,
 RT "BAALC, the human member of a novel mammalian neuroectoderm gene
 RT lineage, is implicated in hematopoiesis and acute leukemia.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
 DR EMBL; AF371326; AAL50522.1; -;
 DR Pfam; PF06989; BAALC N; 1.
 SQ SEQUENCE 54 AA; 5681 MW; 6538C3DABED9B825 CRC64;

Query Match 34.8%; Score 267; DB 2; Length 54;
 Best Local Similarity 92.6%; Pred. No. 8.8e-17;
 Matches 50; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MCGGSRADAIIPRYESWTRETESTWLTYYTDSALPSAAATDSCPEAGGLHAG 54
 DB 1 MCGGSRADAIIPRYESWTRETESTWLTYYTDSADAPPSAAAPDSCPEAGGLHAG 54

RESULT 14
 ID 0801V5 PRELIMINARY; PRT; 32 AA.
 AC 0801V5;
 DT 01-JUN-2003 (TRENBLrel. 24, Created)
 DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
 DE SI: C215113.3 (Novel protein similar to human brain and acute
 DE leukemia, cytoplasmic (BAALC)) (Fragment).
 GN Name=SI: C215113.3; (Zebrafish)
 OS Brachydanio rerio (Zebrafish)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OK NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 RA Sehra H.,
 RL EMBL; AL807244; CAD87801.1; -;
 DR InterPro: IPR009728; BAALC_N.
 DR Pfam; PF06989; BAALC_N; 1.
 FT NON_TER 32
 FT 32
 SQ SEQUENCE 32 AA; 3663 MW; 916A5445D263B7F7 CRC64;

Query Match 21.0%; Score 161; DB 2; Length 32;
 Best Local Similarity 90.6%; Pred. No. 1.7e-07;
 Matches 29; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MCGGSRADAIIPRYESWTRETESTWLTYYTDSALPSAAATDSCPEAGGLHAG 32
 DB 1 MCGGSRADAIIPRYESWTRETESTWLTYYTDSADAPPSAAAPDSCPEAGGLHAG 32

RESULT 15
 ID 06ZMU6 PRELIMINARY; PRT; 613 AA.
 AC 06ZMU6;
 DT 05-JUL-2004 (TRENBLrel. 27, Created)
 DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
 DE FLJ00420 protein (Fragment).
 GN Name=FLJ00420;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OK NCBI_TaxID=9606;
 RN [1]

```

RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Ukuya H., Takano J., Nomura N., Kikuno R., Nagase T., Ohara O.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
RR EMBL: AK160386; BAD18729.1; -.
DR GO: GO:0005634; C:nucleus; IEA.
DR GO: GO:0030528; F:transcription regulator activity; IEA.
DR GO: GO:0045449; P:regulation of transcription; IEA.
DR InterPro: IPR007207; Not N.
DR InterPro: IPR010976; rRNA_binding_arm.
DR Pfam: PF04065; Nct3; 1.
SQ NON_TER 1
SEQUENCE 613 AA; 64243 MW; 8D78E87BD58F9E2C CRC64;

Query Match 12.1%; Score 92.5; DB 2; Length 613;
Best Local Similarity 31.9%; Pred. No. 8;
Matches 30; Conservative 7; Mismatches 32; Indels 25; Gaps 3;

QY 31 TDSALDPSAATDTSGPBAGLH-----ACVLDEGSSNVLPAPAA-- 71
DB 216 TDSVVSQSPANNGSKPHSNCHPOSAPVLPYPSGPPAPASALSTTPNNGVAPAPADS 275
QY 72 --GGIANPEKKMNCCTCCPNQSGLS---SGPLT 99
DB 276 ALGPKASBPAPSHNSGTPAPYAAVAAPAPSGPST 309

Search completed: April 12, 2005, 15:42:08
Job time : 102.667 secs

```

This Page Blank (uspto)

	Query March Best Local Similarity	11.2%; Score 87.5; DB 2;	Length 369; Pred. No.0.36;
	Matches	Conservative 19;	Mismatches 48; Indels 13; Gaps 4;
OY	33 SDAPPAAAPGSGEAGGLHSGMLIEDGLPSNGVPRS--TAPGGIINPEKKTNCETQCPRPQ	91	
Dd	40 SSTSTSSSPFSSFPSSSSSSSCYPILPS--TPBEVSADDETNPPO--SAQTACSSPS	95	
OY	92 SLSSGPL-----TOKONGILOTFEARIDAKRPAKEVTINTVDSIQ	131	
Dd	96 VVALPLPDQSDEGSSSQKEESPFTLQVLDPDESIRPRESIDERKVDLVQ	143	

```

Query Match      11.0%; Score 85.5; DB 4; Length 316;
Best Local Similarity 31.5%; Pred. No. 0.48;
Matches 29; Conservative 12; Mismatches 38; Indels 13; Gaps 5;

QY      63 NGVRS--TAPGGIPNEKKTNC-TCPCNPQSLSSGPLYTKONGLOTTEAKRDA--KR 116
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      76 NFFQSTIDTPAAVAASNPRTNABPAKIPNEKLKTESPLSKONGAATTKEKDVLLET 135
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      117 MPAEVTINTVDSIQGNDR-----SRRTIK 141
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      136 KSTSTTVSNNSNVLYQTLESLPIGVERTIK 167
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 4
US-09-902-540-12404
Sequence 12404, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(115849)B
CURRENT APPLICATION NUMBER: US/09/902,540
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 12404
LENGTH: 656
TYPE: PRT
ORGANISM: Myxococcus xanthus
US-09-902-540-12404

Query Match      10.8%; Score 84; DB 4; Length 656;
Best Local Similarity 26.1%; Pred. No. 2;
Matches 42; Conservative 16; Mismatches 59; Indels 44; Gaps 7;

QY      2 GCCGSRADALEPRYESWTETESTWLTITDSDA-PPSAAPPBG-PBAGGLSGMLEDG 59
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      137 GCDDSEBDK-----TNDCPNPSVGCQGQDSGTFDAGAPDAGNPDAGIPDAGDPDAG 187
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      60 LPSNGVRSTRAP-----GGIPNEKKTNCETO-----CPNP 90
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      188 NPDAHGPDACNPADGPNPDGGGPLPYPTLIESQSASHQVSAGOTVTHVTRGRDDLACPLQ 247
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      91 QS--LSGGPLYTKONGLOTTEAKRDKR-MPAEVTINTV 127
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      248 FSWDASGTLGSPETSATSDVTMTAPHCVLMSSEVITILT 288
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 5
US-09-252-991A-25785
Sequence 25785, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 25785
LENGTH: 208
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
```

US-09-252-991A-25785

Query Match 10.4%; Score 81; DB 4; Length 208;
 Best Local Similarity 23.0%; Pred. No. 0.83;
 Matches 37; Conservative 18; Mismatches 70; Indels 36; Gaps 4;

QY 3 CCGSR-----ADAIEPRYESWTRTESTWLTYSDDAPPSAAAPDGPBAGLHSGMLD 58
 DB 28 CCGSASIRPARAPBPBKPAPSSPTTAMTGSTMTTRSPSTAKARPTTTGASASPATTP 87
 QY 59 GL-----PSNGVRSAPGIPNPEKTCNCTO-----CPV 89
 DB 88 GTTGTPPPAPPTKSSAAGCAMPDCHACRRRTAPRPVPAWISTSRASVPACTC 147
 QY 90 POSTSSGPLTQKONGLOTTAKADAK---RMPAKETINTV 127
 DB 148 PTTSSRTGPTSSRRRTTRCTAPTRCSRYARPAKATTRCT 188

RESULT 6

US-09-949-016-11662
 ; Sequence 11662, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; PRIOR FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: PasteSeq for Windows Version 4.0
 ; SEQ ID NO 11662
 ; LENGTH: 507
 ; TYPE: PRT
 ; ORGANISM: Human
 ; US-09-949-016-11662

Query Match 10.3%; Score 80; DB 4; Length 507;
 Best Local Similarity 29.4%; Pred. No. 3.7;
 Matches 30; Conservative 7; Mismatches 57; Indels 8; Gaps 2;

QY 24 ESTWLTYYTDSAPPSAA-----APDSGEAAGLHSGMLDGLPSNGVRSAPGIPN 76
 DB 181 KSFLLITITVFTNPPOVATYHRAIKTVVDGPREPRRHQKLDKSPSLFSDRLSDLGRIH 240
 QY 77 PEKTCNCTQCPNPOSTSSGPLTQKONGLOTTAKADAKMP 118
 DB 241 PSRWVGVPQNPRP-SLNSAPSPFNPQGSQITDPRQAQSSP 281

RESULT 7

US-09-086-663A-81
 ; Sequence 81, Application US/0908663A
 ; Patent No. 6518063
 ; GENERAL INFORMATION:

; APPLICANT: DUCY, PATRICIA
 ; APPLICANT: KARSSENTY, GERARD
 ; TITLE OF INVENTION: OSF2/CBFA1 COMPOSITIONS AND METHODS OF USE
 ; FILE REFERENCE: UTSC:525
 ; CURRENT APPLICATION NUMBER: US/09/086,663A
 ; PRIOR FILING DATE: 1998-05-29
 ; PRIOR APPLICATION NUMBER: 60/080,189
 ; PRIOR FILING DATE: 1998-03-24
 ; PRIOR APPLICATION NUMBER: 60/048,430
 ; PRIOR FILING DATE: 1997-05-29
 ; NUMBER OF SEQ ID NOS: 83

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 81
 ; LENGTH: 521
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-086-663A-81

Query Match 10.3%; Score 80; DB 4; Length 521;
 Best Local Similarity 29.4%; Pred. No. 3.9;
 Matches 30; Conservative 7; Mismatches 57; Indels 8; Gaps 2;

QY 24 ESTWLTYYTDSAPPSAA-----APDSGEAAGLHSGMLDGLPSNGVRSAPGIPN 76
 DB 195 KSFLLITITVFTNPPOVATYHRAIKTVVDGPREPRRHQKLDKSPSLFSDRLSDLGRIH 254
 QY 77 PEKTCNCTQCPNPOSTSSGPLTQKONGLOTTAKADAKMP 118
 DB 255 PSRWVGVPQNPRP-SLNSAPSPFNPQGSQITDPRQAQSSP 295

RESULT 8

US-09-086-663A-82
 ; Sequence 82, Application US/0908663A
 ; Patent No. 6518063
 ; GENERAL INFORMATION:

; APPLICANT: DUCY, PATRICIA
 ; APPLICANT: KARSSENTY, GERARD
 ; TITLE OF INVENTION: OSF2/CBFA1 COMPOSITIONS AND METHODS OF USE
 ; FILE REFERENCE: UTSC:525
 ; CURRENT APPLICATION NUMBER: US/09/086,663A
 ; PRIOR FILING DATE: 1998-05-29
 ; PRIOR APPLICATION NUMBER: 60/080,189
 ; PRIOR FILING DATE: 1998-03-24
 ; PRIOR APPLICATION NUMBER: 60/048,430
 ; PRIOR FILING DATE: 1997-05-29
 ; NUMBER OF SEQ ID NOS: 83
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 82
 ; LENGTH: 528
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 ; US-09-086-663A-82

Query Match 10.3%; Score 80; DB 4; Length 528;
 Best Local Similarity 29.4%; Pred. No. 3.9;
 Matches 30; Conservative 7; Mismatches 57; Indels 8; Gaps 2;

QY 24 ESTWLTYYTDSAPPSAA-----APDSGEAAGLHSGMLDGLPSNGVRSAPGIPN 76
 DB 202 KSFLLITITVFTNPPOVATYHRAIKTVVDGPREPRRHQKLDKSPSLFSDRLSDLGRIH 261
 QY 77 PEKTCNCTQCPNPOSTSSGPLTQKONGLOTTAKADAKMP 118
 DB 262 PSRWVGVPQNPRP-SLNSAPSPFNPQGSQITDPRQAQSSP 302

RESULT 9

US-09-086-663A-71
 ; Sequence 71, Application US/0908663A
 ; Patent No. 6518063
 ; GENERAL INFORMATION:

; APPLICANT: DUCY, PATRICIA
 ; APPLICANT: KARSSENTY, GERARD
 ; TITLE OF INVENTION: OSF2/CBFA1 COMPOSITIONS AND METHODS OF USE
 ; FILE REFERENCE: UTSC:525
 ; CURRENT APPLICATION NUMBER: US/09/086,663A
 ; PRIOR FILING DATE: 1998-05-29
 ; PRIOR APPLICATION NUMBER: 60/080,189
 ; PRIOR FILING DATE: 1998-03-24
 ; PRIOR APPLICATION NUMBER: 60/048,430
 ; PRIOR FILING DATE: 1997-05-29
 ; NUMBER OF SEQ ID NOS: 83
 ; SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 71
LENGTH: 548
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-086-663A-71

Query Match 10.3%; Score 80; DB 4; Length 548;
Best Local Similarity 29.4%; Pred. No. 4.1; Indels 8; Gaps 2;
Matches 30; Conservative 7; Mismatches 57; Indels 8; Gaps 2;

QY 24 ESTWLTYYTSDAPPSAA-----APDSGPEAGLHSGMLDGLPSNGVPRSTAPGGIN 76
DB 202 KSFLLITVFTNPPQVATYHRAIKTVTDGPREPRRRRQKLDSPKPSLFSDDLSDIGRI 261
QY 77 PEKTNCEIQCPNPQSSGPILOKONGLOTTAKKDAKMP 118
DB 262 PSNRVGVPPQNP RP-SLNSAPSPFPNPGQSQITDPRAQSSP 302

RESULT 10

US-09-086-663A-2
Sequence 2, Application US/09086663A
Patent No. 6518063
GENERAL INFORMATION:
APPLICANT: DUCY, PATRICIA
APPLICANT: KARSENTY, GERARD
TITLE OF INVENTION: OSF2/CBFA1 COMPOSITIONS AND METHODS OF USE
FILE REFERENCE: UTSC:525
CURRENT APPLICATION NUMBER: US/09/086,663A
CURRENT FILING DATE: 1998-05-29
PRIOR APPLICATION NUMBER: 60/080,189
PRIOR FILING DATE: 1998-03-24
PRIOR APPLICATION NUMBER: 60/048,430
PRIOR FILING DATE: 1997-05-29
NUMBER OF SEQ ID NOS: 83
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 596
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-086-663A-2

Query Match 10.3%; Score 80; DB 4; Length 596;
Best Local Similarity 29.4%; Pred. No. 4.7;
Matches 30; Conservative 7; Mismatches 57; Indels 8; Gaps 2;

QY 24 ESTWLTYYTSDAPPSAA-----APDSGPEAGLHSGMLDGLPSNGVPRSTAPGGIN 76
DB 270 KSFLLITVFTNPPQVATYHRAIKTVTDGPREPRRRRQKLDSPKPSLFSDDLSDIGRI 329
QY 77 PEKTNCEIQCPNPQSSGPILOKONGLOTTAKKDAKMP 118
DB 330 PSNRVGVPPQNP RP-SLNSAPSPFPNPGQSQITDPRAQSSP 370

RESULT 11

US-09-086-663A-80
Sequence 80, Application US/09086663A
Patent No. 6518063
GENERAL INFORMATION:
APPLICANT: DUCY, PATRICIA
APPLICANT: KARSENTY, GERARD
TITLE OF INVENTION: OSF2/CBFA1 COMPOSITIONS AND METHODS OF USE
FILE REFERENCE: UTSC:525
CURRENT APPLICATION NUMBER: US/09/086,663A
CURRENT FILING DATE: 1998-05-29
PRIOR APPLICATION NUMBER: 60/080,189
PRIOR FILING DATE: 1998-03-24

PRIOR APPLICATION NUMBER: 60/048,430
PRIOR FILING DATE: 1997-05-29
NUMBER OF SEQ ID NOS: 83
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 80
LENGTH: 596
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-086-663A-80

Query Match 10.3%; Score 80; DB 4; Length 596;
Best Local Similarity 29.4%; Pred. No. 4.7;
Matches 30; Conservative 7; Mismatches 57; Indels 8; Gaps 2;

QY 24 ESTWLTYYTSDAPPSAA-----APDSGPEAGLHSGMLDGLPSNGVPRSTAPGGIN 76
DB 270 KSFLLITVFTNPPQVATYHRAIKTVTDGPREPRRRRQKLDSPKPSLFSDDLSDIGRI 329
QY 77 PEKTNCEIQCPNPQSSGPILOKONGLOTTAKKDAKMP 118
DB 330 PSNRVGVPPQNP RP-SLNSAPSPFPNPGQSQITDPRAQSSP 370

RESULT 12

US-09-902-540-14218
Sequence 14218, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 14218
LENGTH: 1036
TYPE: PRT
ORGANISM: Myxococcus xanthus
US-09-902-540-14218

Query Match 10.3%; Score 80; DB 4; Length 1036;
Best Local Similarity 24.5%; Pred. No. 10;
Matches 27; Conservative 14; Mismatches 47; Indels 22; Gaps 3;

QY 31 TDSAPPSAAAPDSGPEAGLHSGMLDGLPSNGVPRSTAPGGINP-----EKKTNCEI 85
DB 535 SDEASPPPGAAPDHGAP-----PSSESPPANTARGAVPPPLAPQSECEPATEDT 641
QY 86 QCPNPQSSGPILOKONGLOTTAKKDAKMPAKKVIINTVDSIQMDR 135
DB 642 APPSPSPAEDAPT---ALTRLDGPVAPDPDGMELQILSDPEYSPLE 687

RESULT 13

US-09-948-016-7040
Sequence 7040, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYOMPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CLO01307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7040
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7040

Query Match 10.2%; Score 79; DB 4; Length 500;
Best Local Similarity 26.3%; Pred. No. 4.7;
Matches 30; Conservative 14; Mismatches 50; Indels 20; Gaps 4;

QY 9 DAIEPRYESTWRETESTWLTYYTDSAPPSAAPPDSGPEAGLHSGMLEDLPSNGVPRS 68
DB 148 DEAESQYDSCIESLSRL-----RSLPESTASAPASG-----SDGSPQ---PCT 188
QY 69 TAPGIRPEKKTNCETOCNPQISGGLTOKONGLOTTEAKDKAKMPAKY 122
DB 189 HPPGPVXEPQEKEDADGERAD-STYGSSTLYTLILGGEPAEDPAPRLPLPHV 241

RESULT 14

US-09-252-991A-21479
; Sequence 21479, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21479
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21479

Query Match 10.1%; Score 78.5; DB 4; Length 330;
Best Local Similarity 26.2%; Pred. No. 3;
Matches 34; Conservative 13; Mismatches 50; Indels 33; Gaps 6;

QY 2 GCGGSRADAE-----PRYESTWRETESTWLTYYTDSAPPSA----- 39
DB 26 GCMGSDRTVPQLRAGYAKVEQGRILGKQAEASATGLSTVPLPSPMRNLGCAIVT 85
QY 40 -AAPDSGPEAGLHSGMLEDLPSNGVPRSTAPGIRPEKKTNCETOCNPQISG 96
DB 86 DATGGLPRGGGRHRRRRHPSPTHLPRS--PG---KORRLHLRRQAAPAP---AAG 137
QY 97 PLTOKONGLO 106
DB 138 PARQDPGLQ 147

RESULT 15

US-09-976-594-427
; Sequence 427, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS

; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 427
; LENGTH: 462
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 2716815CD1
US-09-976-594-427

Query Match 10.1%; Score 78.5; DB 4; Length 462;
Best Local Similarity 28.6%; Pred. No. 4.7;
Matches 30; Conservative 16; Mismatches 32; Indels 27; Gaps 7;

QY 16 YESWRETESTWLTYYTDSAPPSAAPPDSGPEAGLHSGMLEDLPSNGVPRSTAPGIR 75
DB 298 YRMSAEVTS--TYSDEDRPK--VPPREPLS-----PSNS--RTSPSKSLP 339
QY 76 NPEKKTNCETQ--CPNPQISGGLTOKONGLOTTEAKDKAKMP 118
DB 340 SYLNGVMPPTQSFAPDEKIVSKAL--QRQ-----SEGSASKVP 377

Search completed: April 12, 2005, 15:43:29
Job time : 27.6667 secs

This Page Blank (uspto)


```
A:Molecule type: DNA  
A:Residues: 1-690 <STO>  
A:Cross-references: UNIPROT:Q9FKI7, GB:AEO05172, NID:g10086468, PIDD:AA612528.1, GSPDB:CIGenetics:  
A:Map position: 1
```

```
Query Match      11.4%; Score 88.5; DB 2; Length 690;  
Best Local Similarity 20.5%; Pred. No.7.2;  
Matches    33; Conservative   29; Mismatches     62; Indels   37; Gaps    6;
```

```
OY       16 YESWTRTESTWLVTDSAPP-----SAAAPSGPAGLHSGMLEDGLPSNGVPRST 69  
           |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::||  
Db        24 YDSNTEKYSTT--LETPDPDFPLPVSRRAVIPESEGEGSSCTGFPDASIGNNS---TS 78  
  
OY       70 APGGIFNEPKTNCTCQ-CPNPOSLSGCLPTOKONGLOTRE-----            109  
           ::::>::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
Db        79 SPFGFPAATQASTSESKTKPTSIPTPTPOBSRSTETPMSPVPITQPSPASSIP 138  
  
OY       110 ----AKRDAXMPAKEVTIN--VTDSIQGMRSRRITNCV 144  
           |         |         |         |         |         |         |  
Db        139 TMAATSSASTSQOSSVASNKSTTDVVIOEQASPACTAPCI 179
```

```
RESULT 3  
138659 melanoma antigen Mage-10 - human  
CISpecies: Homo sapiens (man)  
C>Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 09-Jul-2004  
C/Accession: I38659  
R:De Plaen, E.; Arden, K.; Traversari, C.; Gaforio, J.J.; Szikora, J.P.; De Smet, C.; Brice,  
on, T.  
Immunogenetics 40, 360-369, 1994  
A>Title: Structure, chromosomal localization, and expression of 12 genes of the Mage fam  
A:Reference number: I38659; MUID:95012457; PMID:7927540  
A:Accession: I38659  
A>Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-369 <RES>  
A:CROSS-references: UNIPROT:P4J363; EMBL:UI0685; NID:g533510; PIDN:AAA6869.1; PID:g53359  
CIGenetics:  
A:Gene: GDB:MAGEA10; MAGE10  
A:CROSS-references: GDB:331126  
A:Map position: Xq28-Xq28  
A:introns: #status absent  
C:Superfamily: tumor associated protein MAGE
```

```
Query Match      11.2%; Score 87.5; DB 2; Length 369;  
Best Local Similarity 25.9%; Pred. No.4.2;  
Matches    28; Conservative   19; Mismatches     48; Indels   13; Gaps    4;
```

```
OY       33 SDAPPSAAPDGPGAGLGHSMLBDGLPSNGVPRS-TAPGIENPERKTNETOCENPQ 91  
           :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||  
Db        40 SSTGTSSSFPSPPSSSSSSSCVPIIPS--TFEVSADDEFNPQQ--SAQLACGPS 95  
  
OY       92 SLSSGPL-----TGKNGLOTTFAKRDAKKRPKEAVTIINTVDISIQ 131  
           ::|||::|||::|||::|||::|||::|||::|||::|||::|||  
Db        96 VVALPLPDQSDEGSSSOKEESPRTLQVLPLESLPREIDEIKVTDLVQ 143
```

```
RESULT 4  
A96717 unknown protein, 45065-49536 [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Data: 02-Mar-2001 #sequence_reviseion 02-Mar-2001 #text_change 09-Jul-2004  
C/Accession: A96717  
R:Theologis, A.; Eckert, J.R.; Palm, C.U.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, B.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, M.; Hutvazr, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
A.; Li, J.H.; Lin, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Lucos, J.S.; Maitl, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,
```

ker, M.; Mu, D., Yu, G.; Fraser, C.M.; Ventier, J.C.; Davis R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A|Reference number: A86141; MUID:21016719; PMID:11130712
A|Accession: A86717
A|Status: Preliminary
A|Molecule type: DNA
A|Residues: 1-860 <STO>
A|Cross-references: UNIPROT:Q9G982; GB:AEO5173; NID:g7705093; PIDD:AAF67772.1; GSPDB:GN
C|Genetics:
A|Gene: F23010.12
A|Map position: 1

Query Match 11.2%; Score 87.5; DB 2; Length 860;
Best Local Similarity 31.2%; Pred. No. 11;
Matches 29; Conservative 9; Mismatches 36; Indels 19; Gaps 5;

Cy 23 TESTWLVYTSDDAPPSAAAPDSC--PEAGGLHSGMLEBGLPNSNVPNSTAP-GGIPLN---P 77
 | :
Db 98 TASNCLLGGSSSSSGPTTGTGTTGPTTSGPTTGPTTGTGPTTGTGPTTGTGPTTSGPTSGFPNTGP 157
 | :
Cy 78 EKTNCERTCCENPQSLSGFLTKONGLOTTEA 110
 || :
Db 158 NTGTN--TGMPN-----SNGMPTSSS 176

RESULT 5
 A45301
 microtubule-associated protein tau - mouse
 N/Alternate names: microtubule binding protein tau
 C/Species: Mus musculus (house mouse)
 C/Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004
 C/Accession: A45301/ S31658
 R/Couché, D.; Maxwell, C.; Georgieff, I.S.; Liem, R.K.; Shelanski, M.L.; Nunez, J.
 Proc. Natl. Acad. Sci. U.S.A. 89, 4378-4381, 1992
 A/Title: Primary structure of high molecular weight tau present in the peripheral nervous
 A/Reference number: A45301, MUID:92262443, PMID:1374888
 A/Accession: A45301
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-733 <COU>
 A/Cross-references: UNIPROT:P10637
 A/Note: this sequence is inconsistent with the nucleotide translation
 A/Note: sequence extracted from NCBI backbone (NCBIN:102045, NCBI:P102046)
 R/Kenner, U.; Forstner, M.; Huter, H.; Hoeller, G.; Kurzbauer, R.; Zatloukal, K.; Krisp
 submitted to the EMBL data library, May 1992
 A/Description: First observation of mRNA for a tau-protein from murine liver and kidney.
 A/Reference number: S31658
 A/Accession: S31658
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 'T', 529-651 <KEN>
 A/Cross-references: EMBL:Z12133, NID:G54263, PIDN:CAA78121.1, PID:G388534
 C/Superfamily: microtubule-associated protein tau, MAP2/tau repeat homology
 C/Keywords: microtubule binding; tandem repeat
 F:544-574/Domain: MAP2/tau repeat homology <MT1>
 F:575-605/Domain: MAP2/tau repeat homology <MT2>
 F:606-636/Domain: MAP2/tau repeat homology <MT3>
 F:637-668/Domain: MAP2/tau repeat homology <MT4>

```

Query Match      10.9%; Score 84.5; DB 2, Length 733;
Best Local Similarity 26.2%; Pred. No. 17;
Matches 32; Conservative 16; Mismatches 43; Indels 31; Gaps 6

QY      5 GSRDAIE-----PRYESWTR-----TESTWLTYTSDAPPSAAPDS 44
          |||      |||      |||      |||      |||      |||
Db      155 GTRPEDLEKSHPAELLRRGPPQKRGQGQDILSEEEVEDLITVDSSQDPSPSQASLTLP 214
          |||      |||      |||      |||      |||      |||

QY      45 G---PEAGGLHSGWLELDGLPENGVRSTAPGCIENPE---KTCICETCCPNQSLSSGPL 98
          |||      |||      |||      |||      |||      |||
          |||      |||      |||      |||      |||      |||
Db      215 GRAAPQAG--SGSVCGETAS--VPGLPTEBSVPLPADFSKVSALETASQSPRGCGTICPM 269
          |||      |||      |||      |||      |||      |||
          |||      |||      |||      |||      |||      |||
QY      99 TQ 100
          |

```

Db 270 EE 271

RESULT 6

W2MLDP
E2 protein - deer papillomavirus
C:Species: deer papillomavirus
A:Note: host *Odocoileus virginianus* (American white-tailed deer)
C>Date: 26-May-1986 #sequence_revision 28-May-1986 #text_change 24-Feb-1994
C:Accession: A03673
R:Goff, D.E.; Lancaster, W.D.
J. Virol. 56, 85-91, 1985
A:Title: Molecular cloning and nucleotide sequence of deer papillomavirus.
A:Reference number: A93013; MUID:8529225; PMID:2993669
A:Accession: A03673
A:Molecule type: DNA
A:Residues: 1-416 <GRO>
C:Superfamily: papillomavirus E2 protein
C:Keywords: early protein

Query Match 10.8%; Score 84; DB 1; Length 416;
Best Local Similarity 31.7%; Pred. No. 9.6;
Matches 38; Conservative 10; Mismatches 44; Indels 28; Gaps 7;

QY 4 GGSRADAT-----EPRYESWTR-----TESTWLTYSDDAP-PSAAPDSGPEA-- 48
DB 152 GGDADDLFTTMSGRVYELPERDAARSTGTW-TVNDNRTYHSHAPSHRETTIE 210
QY 49 GGLHSGMLEDGLPSNGVPRS--TAPG-----IPNEKKTNCETCCPNQSLSSG 96
DB 211 GLMNSGGRERGRPTSPDAVHLHTPGGNTVHGVRACENRGRSINRPTYSSTPQSPRS 270

RESULT 7

H70699
Probable PPP protein - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: H70699
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulton, J.E.; Taylor, K.; Whitehead, S.; Barrett, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: H70699
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-514 <COL>
A:Cross-references: UNIPROT:P11588; GB:Z60233; GB:AL123456; NID:g3261645; PIDN:CAM02438.
A:Experimental source: strain H37RV
C:Genetic: A:Gene: PPP
F:35-243/Domain: conserved hypothetical protein y100 homology <Y100>

Query Match 10.8%; Score 84; DB 2; Length 514;
Best Local Similarity 25.9%; Pred. No. 12;
Matches 36; Conservative 10; Mismatches 49; Indels 44; Gaps 6;

QY 1 MCGGSRADAIERPRYESWTRTESTWLTYSDDAP-----PSAAPDSGPEAGLH 52
DB 357 MGC-----LSFR-----NELSOISTGSGGPIDCHLMKLEDLRPERAQRAGLP 401
QY 53 SGMLEDC-----LPSNGVPRSTAPGIPNPEKKTNCETCCPNQSLSSGPTLQ 100
DB 402 AGTLDDAIGQLRLAANSLPCCAPAPATSPGRAP--PTSETTEPNTTSSPAP--- 456
QY 101 KONGLOTTEAKRDKMPA 119
DB 457 ---SPTSAAPPTGTTTPA 471

RESULT 8

I40333
tracheal colonization factor A precursor - Bordetella pertussis
N:Alternate names: tcfa protein
C:Species: Bordetella pertussis
C>Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 09-Jul-2004
C:Accession: S77633; I40333
R:Pinn, T.M.; Stevens, D.A.
Mol. Microbiol. 16, 625-634, 1995
A:Title: Tracheal colonization factor: a Bordetella pertussis secreted virulence determ
A:Reference number: I40333; MUID:96065692; PMID:7476158
A:Accession: S77633
A:Molecule type: DNA
A:Residues: 1-672 <PIN>
A:Cross-references: UNIPROT:Q45343; EMBL:U16754; NID:g984282; PIDN:AA043453.1; PID:g984.
A:Experimental source: strain 18323
C:Genetic: A:Gene: tcfa
F:1-39/Domain: signal sequence #status predicted <SIG>
F:40-672/Product: tracheal colonization factor A #status predicted <MAT>

Query Match 10.8%; Score 84; DB 2; Length 672;
Best Local Similarity 31.2%; Pred. No. 17;
Matches 35; Conservative 11; Mismatches 44; Indels 22; Gaps 6;

QY 30 YTDSDAP-----SAAAPDSGPEAGLHSGMLEDGLPSNGVPRSTA--PGGI-----PNPE 78
DB 103 HKDNPSPPVVGVPGMASSSGCHNPVGCGTHENGILPGIGKVGSAAPGGLGRNDENSE 162
QY 79 KKTNCETCCPNP---OSLSSGPTLKONGLOTTE-----AKRDARMPAKE 121
DB 163 SSLNPTLGPSPDPDTSTSGSPDCAAGAGSTSPASGAGADA--MEPSE 212

RESULT 9

S49619
crta protein - Rhodobacter sphaeroides
C:Species: Rhodobacter sphaeroides
C>Date: 05-Mar-1995 #sequence_revision 12-May-1995 #text_change 09-Jul-2004
C:Accession: S49619
R:Yang, H.P.; Cogdell, R.D.; Takachi, S.; Hunter, C.N.
submitted to the EMBL Data Library, November 1994
A:Description: The complete DNA sequence, specific TMS insertion map and gene assignment.
A:Reference number: S49619
A:Accession: S49619
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-327 <LAN>
A:Cross-references: UNIPROT:Q54790; EMBL:X82458; NID:g575405; PID:g575406
C:Genetic: A:Gene: crta
C:Superfamily: spheroidene monooxygenase

Query Match 10.7%; Score 83.5; DB 2; Length 327;
Best Local Similarity 23.9%; Pred. No. 8;
Matches 33; Conservative 15; Mismatches 49; Indels 41; Gaps 5;

QY 2 GCGSRADAIERPRYESWTR-----TESTW-----LTYTDSAPPSAAP 42
DB 190 GAGGPHGEAIVKVRANWFKELVYARFOLLGIGKEGDPVGEALTAIPSEAPKAPAP 249
QY 43 DSGPEAGLHSGMLEDGLPSNGVPRSTAPGIPNPEKKTNCETCCPN---PQSLSSGP 97
DB 250 AAAQPA-----PAAEAPKAPAPAAVEKPAALAVEMKPAEPKPVVEAP 292
QY 98 LTKONGLOTTEAKRDK 115
DB 293 KPFKSKGLEADNAGREAR 310

RESULT 10

IJMSNL
neural cell adhesion molecule 1 precursor, long domain splice form - mouse

N:Alternate names: NCAM-180
N:Contains: neural cell adhesion molecule, short domain splice form (NCAM-140)
C:Species: Mus musculus (house mouse)
C:Date: 31-Mar-1993 #sequence,revision 31-Mar-1993 #text_change 09-Jul-2004
C:Accession: A29673; S00844; S00384; A28281; A44290; S00383
R:Barthele, D.; Santoni, M.J.; Wille, W.; Ruppert, C.; Chaux, J.C.; Hirsch, M.R.; Fontec
EMBO J. 6, 907-914, 1987
A:Title: Isolation and nucleotide sequence of mouse NCAM cDNA that codes for a Mr 79,000
A:Reference number: A29673; MUID:87246524; PMID:3595563
A:Accession: A29673
A:Molecule type: mRNA
A:Residues: 1-548; 'T', 550-571, 'T', 573-574, 'D', 576-588, 'MPS', 593, 'S', 595-599, 'P', 601, 'L'
A:Cross-references: UNIPROT:P13595; EMBL:Y00051; NID:G53347; PIDD:CA68263.1; PID:G53343
R:Sanon, M.J.; Barthele, D.; Barba, J.A.; Hirsch, M.R.; Steinmetz, M.; Goridis, C.; W
Nucleic Acids Res. 15, 8621-8641, 1987
A:Title: Analysis of cDNA clones that code for the transmembrane forms of the mouse neur
A:Reference number: S00844; MUID:88067687; PMID:3684567
A:Accession: S00844
A:Molecule type: mRNA
A:Residues: 529-809, 1077-1115 <SN>
A:Cross-references: EMBL:X06328; NID:G53322; PIDD:CAA29641.1; PID:G817984
R:Barba, J.A.; Chaux, J.C.; Steinmetz, M.; Goridis, C.
EMBO J. 7, 625-632, 1988
A:Title: Differential splicing and alternative polyadenylation generates distinct NCAM t
A:Reference number: S00382; MUID:88283628; PMID:3396534
A:Accession: S00384
A:Molecule type: DNA
A:Residues: 642-1115 <BAR>
A:Cross-references: EMBL:X07195
R:Barthele, D.; Vopper, G.; Wille, W.
Nucleic Acids Res. 16, 4217-4225, 1988
A:Title: NCAM-180, the large isoform of the neural cell adhesion molecule of the mouse,
A:Reference number: A28281; MUID:88247737; PMID:2454455
A:Accession: A28281
A:Molecule type: mRNA
A:Residues: 804-1081 <BA3>
A:Cross-references: EMBL:X07244; NID:G53321; PIDD:CAA30230.1; PID:G929720
R:Rougou, G.; Marshak, D.R.
J. Biol. Chem. 261, 3396-3401, 1986
A:Title: Structural and immunological characterization of the amino-terminal domain of n
A:Reference number: A44290; MUID:86140120; PMID:3512556
A:Accession: A44290
A:Molecule type: protein
A:Residues: 20-36 <ROU>
C:Comment: NCAM mediates cell-cell adhesion via homophilic binding with another NCAM mol
C:Comment: NCAM mediates cell-cell adhesion via homophilic binding with another NCAM mol
C:Gene: NCAM
C:Genetics:
A:Map position: 9
A:Intron: 643/3; 701/1; 770/2; 809/2; 1076/2
C:Superfamily: neural cell adhesion molecule; fibronectin type III repeat homology; immu
C:Keywords: alternative splicing; brain; cell adhesion; duplication; heparin binding; si
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-115/Product: neural cell adhesion molecule, long domain splice form #status experi
F:809-809,1077-1115/Product: neural cell adhesion molecule, short domain splice form #sta
F:120-711/Domain: extracellular #status predicted <EXT>
F:20-711/Domain: immunoglobulin homology <IMM>
F:34-98/Domain: immunoglobulin homology <IMM2>
F:134-191/Domain: immunoglobulin homology <IMM3>
F:152-156/Region: heparin binding #status predicted
F:161-165/Region: heparin binding #status predicted
F:228-290/Domain: immunoglobulin homology <IMM3>
F:262-272/Region: NCAM binding #status predicted
F:323-388/Domain: immunoglobulin homology <IMM4>
F:420-482/Domain: immunoglobulin homology <IMM5>
F:519-596/Domain: fibronectin type III repeat homology <FN3>
F:625-688/Domain: fibronectin type III repeat homology <FN3B>
F:712-729/Domain: transmembrane #status predicted <TM>
F:730-1115/Domain: intracellular #status predicted <INT>
F:41-96,139-189,235-288,330-386,427-480/Diisulfide bonds: #status predicted
F:222,316,348,424,450,479/Binding site: carbohydrate (asn) #status predicted

```

Matches      35; Conservative   15; Mismatches    59; Indels     23; Gaps       6;

QY          6 SRADAIIPRYESWTRETESITWLTYSDDAPSPSAAAPDSCPEAGLHSGMLEDLPSNGV 65
            : ::::|::::::::::|||::|::|::|::|::|::|::|::|::|::|::|::|::|
DB           830 TNSDTTETEFATIAONSPTSET-TTLTSIIAPPATTVPDSNVPAG-----CATPSKGV 881
               |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY          66 -PRSTAGGTPNEP---KTNCENCGNPQSLS-----GPLTKONGLOTTEAK 111
            : ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB           882 TASSSSPSAKAVALVDLDTPFTSPASASNLSSTVLANCAVLSPSPASAC-ETS KAP 940
                |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY          112 RDAKMMPAKEVT 123
             :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB           941 PASKASPAPTPT 952


RESULT 11
T35985
Probable large Pro/Ala/Gly-rich protein - Streptomyces coelicolor
C.Species: Streptomyces coelicolor
C.Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C.Accession: J75985
R.Murphy, L.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A. submitted to the EMBL Data Library, February 1999
A.Reference number: Z21589
A.Accession: J75985
A.Status: Preliminary; translated from GB/EMBL/DDBBJ
A.Molecule type: DNA
A.Residues: 1-1366 <MUR>
A.Cross-references: UNIRROT.Q9Z529; EMBL.AL035559; PIDN.CAB37473.1; GSPPDB:GN00070; SCOPED A.Experimental source: strain A3(2)
A.GeneticB:
A.Gene: SCOEBD:SCOFP2_06c


Query Match              10.6%; Score 82.5; DB 2; Length 1366;
Best Local Similarity 30.8%; Pred. No. 51;
Matches 33; Conservative 12; Mismatches 25; Indels 37; Gaps 8;

QY         2 GGGSRRDAIIPRYESWTRETESTWLTYTSDAPSPAAPD--SCP----EAGGLHG- 53
            |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB          458 GAGAAR-----PTMAKEAPS-----PPASTAPEWPMSGPGAVGCAGVS SP 499
               |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY        54 --GMLEDG-LPSNVPR-----STARGGINPEKKTNCTENCPCNP 90
            |:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB          500 GTGLRGSGTVAAAGVPLFGARGVGGAAGVPLTFVTASERP-PGP 545


RESULT 12
S22571
Integrase-like protein FE65 - rat
C.Species: Rattus norvegicus (Norway rat)
C.Date: 31-Dec-1993 #sequence_revision 02-Aug-1994 #text_change 09-Jun-2004
C.Accession: S22571; S22572
R.Bullito, A.; Zambrano, N.; Mogavero, A.R.; Amendola, R.; Cimino, F.; Russo, T. Nucleic Acids Res. 19, 5269-5274, 1991
A.Title: A rat brain mRNA encoding a transcriptional activator homologous to the DNA bin
A.Reference number: S22571; MUID:92020215; PMID:1923810
A.Accession: S22571
A.Molecule type: mRNA
A.Residues: 1-499 <DUU1>
A.Cross-references: UNIRROT.Q09MK3; EMBL.X60465; NID:g57559; PIDN:CMA42999.1; PID:g57560
A.Accession: S22572
A>Status: nucleic acid sequence not shown; translation not shown
A.Molecule type: DNA
A.Residues: 31-318 <DU12>
A.Cross-references: EMBL:X60468; NID:g57561; PIDN:CMA42998.1; PID:g1177617
A.Note: This sequence was submitted to the EMBL Data Library, July 1991
C.Genetics:
C.Introns: 88/3; 107/3; 135/3; 157/3; 207/3; 250/2; 252/2; 290/3
C.Keywords: transcription regulation
F:42-78/DNAin: WW repeat homology <WW1>


Query Match              10.5%; Score 81.5; DB 2; Length 499;
Best Local Similarity 25.3%; Pred. No. 19;
```


this Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 15:16:23 ; Search time 94.6667 Seconds

(without alignments)
784.347 Million cell updates/sec

Title: US-10-705-716A-4

Perfect score: 778

Sequence: 1 MCGCGSRADAIERYESWT.....VTDSIQMDRSRRITKNCVN 145

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	778	100.0	145	2 Q9HA93	Q9HA93 homo sapien
2	750.5	96.5	180	2 Q8WA93	Q8WA93 homo sapien
3	664	85.3	145	2 Q8WNE9	Q8WNE9 mus scrofa
4	649	83.4	145	2 Q8VH11	Q8VH11 mus musculu
5	645	82.9	145	2 Q920K5	Q920K5 rattus norv
6	569.5	73.2	149	2 Q8WXS1	Q8WXS1 homo sapien
7	381	49.0	123	2 Q9CYS9	Q9CYS9 mus musculu
8	309	39.7	73	2 Q8WXS0	Q8WXS0 homo sapien
9	299	38.4	54	2 Q8WTP6	Q8WTP6 homo sapien
10	293	37.7	80	2 Q8WXS2	Q8WXS2 homo sapien
11	282	36.2	54	2 Q8WNE8	Q8WNE8 mus scrofa
12	278	35.7	54	2 Q8VBS8	Q8VBS8 mus musculu
13	278	35.7	54	2 Q790N3	Q790N3 rattus norv
14	161	20.7	32	2 Q801V5	Q801V5 brachydanio
15	97.5	12.5	306	2 Q72H02	Q72H02 thermus the
16	95.5	12.3	1097	1 S24C ARATH	Q9M291 arabidopsis
17	94.5	12.1	1625	2 O55597	O55597 garlic vtru
18	93	12.0	291	2 Q6C2R4	Q6C2R4 yarrowia li
19	92.5	11.9	717	2 Q6FKP2	Q6FKP2 candida gla
20	92.5	11.9	1001	1 P5PA RAT	Q9JME1 rattus norv
21	92	11.8	221	2 Q8GAN3	Q8GAN3 archibacte
22	91.5	11.8	222	2 Q8LE41	Q8LE41 arabidopsis
23	91.5	11.8	249	2 Q9ZXR5	Q9ZXR5 bacterioph
24	91.5	11.8	2042	2 Q767L8	Q767L8 sus scrofa
25	91	11.7	189	2 Q73W33	Q73W33 mycobacteri
26	90.5	11.6	670	2 Q828T2	Q828T2 streptomyce
27	90.5	11.6	330	2 Q6BXP9	Q6BXP9 debrayomyce
28	90	11.6	219	2 Q82AV5	Q82AV5 streptomyce
29	89	11.4	321	2 Q9VUD8	Q9VUD8 drosophila
30	89	11.4	629	2 Q87V71	Q87V71 pseudomonas
31	89	11.4	805	2 Q758W4	Q758W4 ashyia goss

32	89	11.4	1064	2 Q8DIPI	Q8DIPI synechococ
33	88.5	11.4	340	2 Q96IZ0	Q96IZ0 homo sapien
34	88.5	11.4	690	2 Q9FX17	Q9FX17 arabidopsis
35	88	11.3	302	2 Q93Z47	Q93Z47 arabidopsis
36	88	11.3	472	2 Q8UVG3	Q8UVG3 gallus gall
37	88	11.3	472	2 Q8UWC8	Q8UWC8 gallus gall
38	88	11.3	1194	2 Q6C4Z0	Q6C4Z0 yarrowia li
39	87.5	11.2	192	2 Q67ZP7	Q67ZP7 arabidopsis
40	87.5	11.2	222	2 Q93V72	Q93V72 arabidopsis
41	87.5	11.2	340	2 Q6FHY9	Q6FHY9 homo sapien
42	87.5	11.2	369	1 MAGA HUMAN	P43363 homo sapien
43	87.5	11.2	754	2 Q7TIP0	Q7TIP0 brachydanio
44	87.5	11.2	758	2 Q7SDA2	Q7SDA2 neurospora
45	87.5	11.2	860	2 Q9C982	Q9C982 arabidopsis

ALIGNMENTS

RESULT 1	ID	Q9HA93	PRELIMINARY;	PRT;	145 AA.
AC	Q9HA93				
DT	01-MAR-2001 (T-REMBLrel. 16, Created)				
DT	01-MAR-2001 (T-REMBLrel. 16, Last sequence update)				
DT	25-OCT-2004 (T-REMBLrel. 28, Last annotation update)				
DE	Hypochemical protein FLJ12015 (BAAIC isoform 1-6-8) (Brain and acute				
DE	Leukemia, cytoplasmic) (BAAIC 1-6-8).				
GN	Name=BAAIC;				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_TaxId=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Whole embryo;				
RX	PubMed=14702039; DOI=10.1038/ng1285;				
RA	Oca T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,				
RA	Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,				
RA	Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,				
RA	Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K.,				
RA	Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,				
RA	Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,				
RA	Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,				
RA	Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,				
RA	Nimomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,				
RA	Tanai H., Kimata M., Watanabe M., Hirooka S., Chiba Y., Ishida S.,				
RA	Ono Y., Takiguchi S., Watanabe S., Yoshida M., Hoshino J., Kusano J.,				
RA	Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,				
RA	Togiyama S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,				
RA	Musaehino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,				
RA	Yoshikawa Y., Matsumura H., Ichihara T., Shihata N., Sano S.,				
RA	Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,				
RA	Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,				
RA	Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,				
RA	Yamagaki M., Watanabe K., Kamagai A., Itakura S., Fukunumi Y.,				
RA	Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,				
RA	Ono T., Yamada K., Fujii Y., Ozaki K., Hiraio M., Okmori Y.,				
RA	Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikemura Y., Okamoto S.,				
RA	Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,				
RA	Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,				
RA	Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,				
RA	Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,				
RA	Okumura K., Nagase T., Nomura N., Kikkawa H., Maeuho Y., Yamashita R.,				
RA	Nakai K., Yada T., Nakamura Y., Ohara O., Isozaki T., Sugano S.,				
RT	"Complete sequencing and characterization of 21,243 full-length human				
RT	CDNAs."				
RL	Nat. Genet. 36:40-45(2004).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;				
RA	Tanner S.M., Austin J.L., Leone G., Rush L.J., Plase C., Helonen K.,				
RA	Mrozek K., Sill H., Knuutila S., Kolitz J.E., Archer K.J.,				

```

RA Calliguri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BALC, the human member of a novel mammalian neuroectoderm gene
RT lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
RN [3]
RN SEQUENCE FROM N.A.
RC TISSUE=Brain;
RC MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stropstein M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loughellano N.A., Peters G.J., Abramson R.D., Miliady S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelem M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.C.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalhus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Matra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RN SEQUENCE FROM N.A.
RC TISSUE=Brain;
RC Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK020277; BAB13960.1; -
DR EMBL; AF371319; AAL50515.1; -
DR EMBL; BC011517; AAL1517.1; -
DR EMBL; AF363578; AAL50377.1; -
DR InterPro; IPR009728; BALC_N.
DR Pfam; PF06989; BALC_N; 1.
SQ SEQUENCE 145 AA; 1551 MW; CFB92BBE83DD92E CRC64;

Query Match 100.0%; Score 778; DB 2; Length 145;
Best Local Similarity 100.0%; Pred. No. 3, 7e-55;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGGSRADAIIPRYESWTRETSWTLYTDSADAPPSAAAPSGEAGLHSGMLEDGL 60
DB 1 MCGGSRADAIIPRYESWTRETSWTLYTDSADAPPSAAAPSGEAGLHSGMLEDGL 60
QY 61 PSNGVPRSTAPGGINPEKKTNCETOCNPQSLSSGPILOKONGLOTTEAKRDARMPAK 120
DB 61 PSNGVPRSTAPGGINPEKKTNCETOCNPQSLSSGPILOKONGLOTTEAKRDARMPAK 120
QY 121 EVTINVTDSIQMDRSRRITKNCVN 145
DB 121 EVTINVTDSIQMDRSRRITKNCVN 145

RESULT 2
Q8WXS3 PRELIMINARY; PRT; 180 AA.
AC Q8WXS3;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE BALC 1-5-6-8.
OS Name=BALC; GN
OS Homo sapiens (Human).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tamer S.M., Austin J.L., Leone G., Rush L.J., Plass C., Heinonen K.,
RA Mroczek K., Sill H., Knutti S., Koltz J.E., Archer K.J.,
RA Calliguri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BALC, the human member of a novel mammalian neuroectoderm gene
RT lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AF363578; AAL50379.1; -
DR Pfam; PF06989; BALC_N; 1.
SQ SEQUENCE 180 AA; 19224 MW; 380183E0F188F684 CRC64;

Query Match 96.5%; Score 750.5; DB 2; Length 180;
Best Local Similarity 80.6%; Pred. No. 7, 8e-53;
Matches 145; Conservative 0; Mismatches 0; Indels 35; Gaps 1;

QY 1 MCGGSRADAIIPRYESWTRETSWTLYTDSADAPPSAAAPSGEAGLHSGMLEDGL 53
DB 1 MCGGSRADAIIPRYESWTRETSWTLYTDSADAPPSAAAPSGEAGLHSGMLEDGL 60
QY 54 -----GMLDGEPSNGVPRSTAPGGINPEKKTNCET 85
DB 61 KKAAPTDSVSDGLFSASMAPLAVPSHGMLEDGLPSNGVPRSTAPGGINPEKKTNCET 120
QY 86 QCPNPQSLSSGPILOKONGLOTTEAKRDARMPAKVITINVTDSIQMDRSRRITKNCVN 145
DB 121 QCPNPQSLSSGPILOKONGLOTTEAKRDARMPAKVITINVTDSIQMDRSRRITKNCVN 180

RESULT 3
Q8WNE9 PRELIMINARY; PRT; 145 AA.
AC Q8WNE9;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 26, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE BALC isoform 1-6-8.
GN Name=BALC;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxId=9823;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tamer S.M., Austin J.L., Leone G., Rush L.J., Plass C., Heinonen K.,
RA Mroczek K., Sill H., Knutti S., Koltz J.E., Archer K.J.,
RA Calliguri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BALC, the human member of a novel mammalian neuroectoderm gene
RT lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AF371322; AAL50518.1; -
DR Pfam; PF06989; BALC_N; 1.
SQ SEQUENCE 145 AA; 15401 MW; C75ED7D00BF82E26 CRC64;

Query Match 85.3%; Score 664; DB 2; Length 145;
Best Local Similarity 84.1%; Pred. No. 5, 6e-46;
Matches 122; Conservative 12; Mismatches 11; Indels 0; Gaps 0;

QY 1 MCGGSRADAIIPRYESWTRETSWTLYTDSADAPPSAAAPSGEAGLHSGMLEDGL 60
DB 1 MCGGSRADAIIPRYESWTRETSWTLYTDSADAPPSAAAPSGEAGLHSGMLEDGL 60
QY 61 PSNGVPRSTAPGGINPEKKTNCETOCNPQSLSSGPILOKONGLOTTEAKRDARMPAK 120
DB 61 PSNGVPRSTAPGGINPEKKTNCETOCNPQSLSSGPILOKONGLOTTEAKRDARMPAK 120
QY 121 EVTINVTDSIQMDRSRRITKNCVN 145
DB 121 EVTINVTDSIQMDRSRRITKNCVN 145

RESULT 4
Q8VHVL

```

ID Q8VHV1 PRELIMINARY; PRT; 145 AA.
 AC Q8VHV1;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE BAAIC isoform 1-6-8 (Mus musculus 16 days neonate cerebellum cDNA,
 DE RIKEN full-length cDNA library, clone:963002H16 product:brain and
 DE acute leukemia, cytoplasmic, full insert sequence).
 GN Name=BaalC;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=J2986/SEV7AC;
 RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
 RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plas C., Heinoen K.,
 RA Mrozek K., Sill H., Knuttila S., Koltz J.E., Archer K.J.,
 RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
 RT "BaalC, the human member of a novel mammalian neuroectoderm gene
 RT lineage, is implicated in hematopoiesis and acute leukemia.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
 RX MEDLINE=99279233; PubMed=10349686; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Meth. Enzymol. 303:19-44(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
 RX The PANTOM Consortium;
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Komori H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,
 RA Komori H., Akiyama J., Nishi K., Katsunari T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwake S., Inoue K., Togawa Y., Iizawa M., Ohtsuka E., Watanabe K.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuda S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 RN [7]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
 RX Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,

RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kaga I., Kasukawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Komori H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akehira S., Takeda Y., Tanaka T.,
 RA Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
 RL Submitted (Apr-2002) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF371320; AAL50516.1; -;
 DR EMBL; AK079337; BAC37611.1; -;
 DR MGI; MGI:1928704; BaalC.
 DR Pfam; PF06989; BaalC_N; 1.
 SQ SEQUENCE 145 AA; 15515 MW; 4972670A618C4D6D CRC64;
 QY
 Query Match 83.4%; Score 649; DB 2; Length 145;
 Best Local Similarity 83.4%; Pred. No. 9.1e-45;
 Matches 121; Conservative 10; Mismatches 14; Indels 0; Gaps 0;
 Db 1 MCCGSRADAIIPRYESWTRETESTTWLTYSDAAPSAADSGPEAGLHSGMLDGL 60
 1 MCCGSRADAIIPRYESWTRETESTTWLTYSDAAPSAADSGPEAGLHSGMLDGL 60
 Db 61 PANGVPRSTAPGCIENPEKTCNCECTQCPNPOSLSGPTLQKONGLOTTEAKRDKMPAK 120
 61 SNGVLRPAAPGCIANPEKKNCGTCQPSQNSLSSGPTLQKONGLOTTEAKRDKMSAR 120
 QY 121 EVTINVTDSIQMDRSRRITKNCVN 145
 121 EVTINVTDSIQMDRSRRITKNCVN 145
 Db 121 EVAINVTENIRQMDRSKRTKNCIN 145
 121 EVAINVTENIRQMDRSKRTKNCIN 145
 RESULT 5
 ID Q920K5 PRELIMINARY; PRT; 145 AA.
 AC Q920K5;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Dem-A20-4 (BaalC isoform 1-6-8).
 GN Name=Dem-A20-4; Synonyms=BaalC;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wang X., Tian Q., Li W., Okano A., Suzuki T.;
 RL Submitted (Oct-2001) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley;
 RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
 RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plas C., Heinoen K.,
 RA Mrozek K., Sill H., Knuttila S., Koltz J.E., Archer K.J.,
 RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
 RT "BaalC, the human member of a novel mammalian neuroectoderm gene
 RT lineage, is implicated in hematopoiesis and acute leukemia.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
 DR EMBL; AB073318; BAB70507.1; -;
 DR EMBL; AF371321; AAL50517.1; -;
 DR Pfam; PF06989; BaalC_N; 1.
 SQ SEQUENCE 145 AA; 15475 MW; D5A27AD67456F341 CRC64;
 QY
 Query Match 82.9%; Score 645; DB 2; Length 145;
 Best Local Similarity 83.4%; Pred. No. 1.9e-44;
 Matches 121; Conservative 10; Mismatches 14; Indels 0; Gaps 0;
 Db 1 MCCGSRADAIIPRYESWTRETESTTWLTYSDAAPSAADSGPEAGLHSGMLDGL 60
 1 MCCGSRADAIIPRYESWTRETESTTWLTYSDAAPSAADSGPEAGLHSGMLDGL 60

```
OY 61 PSNGVPRSTAPGGIPNPEKKTNCETCCPNQSSSGPLTQKONGLOTTEAKRDAKMPAK 120
DB 61 SSNGVLRPAAPGGIAPEKKMNCCTCCPNQSSSGPLTQKONGLOTTEAKRDAKMSAR 120
OY 121 EYINTVTDSTQCDRSPRTKNCVN 145
DB 121 EVAISVTENIRQMDRSKRYTKNCIN 145

RESULT 6
O8KXSI PRELIMINARY; PRT; 149 AA.
AC O8KXSI;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE BAALC 1-5-6-7-8.
GN Name=BAALC;
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plass C., Heimonen K.,
RA Mrozek K., Sill H., Knuttila S., Kolitz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BAALC, the human member of a novel mammalian neuroectoderm gene
RT lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AF363578; AAL50381.1; -.
DR Genev; HGNC:14333; BAALC.
DR Pfam; PF06989; BAALC.N:1.
SQ SEQUENCE 149 AA; 15577 MW; DB26C19969B91F6E CRC64;

Query Match 73.2%; Score 569.5; DB 2; Length 149;
Best Local Similarity 75.7%; Pred. No. 2.4e-38;
Matches 109; Conservative 0; Mismatches 0; Indels 35; Gaps 1;

OY 1 MCGGSGRADAIERYRYESTRTRESTMWLTYYTSDAPPSAAPPSSGEGAGLHS----- 53
DB 1 MCGGSGRADAIERYRYESTRTRESTMWLTYYTSDAPPSAAPPSSGEGAGLHSVLEAKS 60
OY 54 -----GMLDGLPSNGVPRSTAPGGIPNPEKKTNCET 85
DB 61 KIKAPDVSDEGLFSAKMAPIAVFSKGLDGLPSNGVPRSTAPGGIPNPEKKTNCET 120
OY 86 QCPNPQSSSGPLTQKONGLOTTE 109
DB 121 QCPNPQSSSGPLTQKONGLOTTE 144

RESULT 7
O9CY99 PRELIMINARY; PRT; 123 AA.
AC O9CY99;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length
DE enriched library, clone:2810457D07 product:brain and acute leukemia,
DE cytoplasmic, full insert sequence.
GN Name=BaalC;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=whole body;
DR MEDLINE=93279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
```

```
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=whole body;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=whole body;
RA The FANTOM Consortium;
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=whole body;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.,
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=whole body;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Kitsuana T., Tashiro H., Itoh M.,
RA Sumi N., Iehi Y., Nakamura S., Hazama M., Nishine T., Harada S.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanuki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=whole body;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hangaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imomaki K., Iehi Y., Itoh M., Izawa M., Kakuwa T., Kato H.,
RA Kawai J., Kojima Y., Kono H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shitaki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK013358; BAB28808.1; -.
DR MGD; MGI:1928704; BaalC.
SQ SEQUENCE 123 AA; 13391 MW; 33DFCB5F8CFA0AB CRC64;

Query Match 49.0%; Score 381; DB 2; Length 123;
Best Local Similarity 77.4%; Pred. No. 2.9e-23;
Matches 72; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

OY 53 SGMLEDGIPSNVPRSTAPGGIPNPEKKTNCETCCPNQSSSGPLTQKONGLOTTEAKR 112
DB 31 SGVLEDGSSNGVLRPAAPGGIAPEKKMNCCTCCPNQSSSGPLTQKONGLOTTEAKR 90
OY 113 DAKRMPAKEVTINTVTDSTQCDRSPRTKNCVN 145
DB 91 DAKRMSAREVAINTENIRQMDRSKRYTKNCIN 123

RESULT 8
```

```

Q8MXS0
ID O8MXS0 PRELIMINARY; PRT; 73 AA.
AC O8MXS0;
DT 01-MAR-2002 (TEMBLrel. 20, Created)
DT 01-MAR-2002 (TEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE BAALC 1-4-5-6-8.
Name=BAALC;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plase C., Heinonen K.,
RA Mrozek K., Sill H., Knuttila S., Koltz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BAALC, the human member of a novel mammalian neuroectoderm gene
RT lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
SQ SEQUENCE 73 AA; 7871 MW; 98DBC28E6EF524A CRC64;

Query Match 39.7%; Score 309; DB 2; Length 73;
Best Local Similarity 96.6%; Pred. No. 9.8e-18;
Matches 56; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MCGGSRADAIEPRYESTRTTESTLTYTDSAPPSAAPPDGPAGLHSGMED 58
Db 1 MCGGSRADAIEPRYESTRTTESTLTYTDSAPPSAAPPDGPAGLHSGCLEF 58

RESULT 9
Q8WTP6 PRELIMINARY; PRT; 54 AA.
AC O8WTP6;
DT 01-MAR-2002 (TEMBLrel. 20, Created)
DT 01-MAR-2002 (TEMBLrel. 20, Last sequence update)
DT 05-JUL-2004 (TEMBLrel. 27, Last annotation update)
DE BAALC isoform 1-8 (BAALC protein) (BAALC 1-8).
Name=BAALC;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plase C., Heinonen K.,
RA Mrozek K., Sill H., Knuttila S., Koltz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BAALC, the human member of a novel mammalian neuroectoderm gene
RT lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
SQ SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Krausberg R.L., Feinold E.A., Grouse L.H., Derge J.G.,
RA Krausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uddin T.B., Toohilyki S., Cantucci P., Prange C.,
RA Bosa S.S., Loguelli N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richarde S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

```

```

RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schmeck A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF371323; AAL50519.1; -
DR EMBL: BC035038; AAL35038.1; -
DR EMBL: AF363578; AAL50378.1; -
DR Pfam: PF06989; BAALC_N; 1.
SQ SEQUENCE 54 AA; 5663 MW; FEF084EABED9B829 CRC64;

Query Match 38.4%; Score 299; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 4.4e-17;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGGSRADAIEPRYESTRTTESTLTYTDSAPPSAAPPDGPAGLHSG 54
Db 1 MCGGSRADAIEPRYESTRTTESTLTYTDSAPPSAAPPDGPAGLHSG 54

RESULT 10
Q8MXS2 PRELIMINARY; PRT; 80 AA.
AC O8MXS2;
DT 01-MAR-2002 (TEMBLrel. 20, Created)
DT 01-MAR-2002 (TEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE BAALC 1-2.
Name=BAALC;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plase C., Heinonen K.,
RA Mrozek K., Sill H., Knuttila S., Koltz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BAALC, the human member of a novel mammalian neuroectoderm gene
RT lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL: AF363578; AAL50382.1; -
DR Pfam: PF06989; BAALC_N; 1.
SQ SEQUENCE 80 AA; 8573 MW; 70BE9B91C3245E2D CRC64;

Query Match 37.7%; Score 293; DB 2; Length 80;
Best Local Similarity 100.0%; Pred. No. 2.1e-16;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGGSRADAIEPRYESTRTTESTLTYTDSAPPSAAPPDGPAGLHSG 53
Db 1 MCGGSRADAIEPRYESTRTTESTLTYTDSAPPSAAPPDGPAGLHSG 53

RESULT 11
Q8WNE8 PRELIMINARY; PRT; 54 AA.
AC O8WNE8;
DT 01-MAR-2002 (TEMBLrel. 20, Created)
DT 01-MAR-2002 (TEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE BAALC isoform 1-8.
Name=BAALC;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin U.L., Leone G., Rush U.J., Plass C., Heimonen K.,
RA Mrozek K., Sill H., Knuttila S., Koltz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BALC, the human member of a novel mammalian neuroectoderm gene
RT lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AF371326; AAL50522.1; -.
DR Pfam; PF06989; BAALC_N; 1.
SQ SEQUENCE 54 AA; 5667 MW; 6538C3DABED9B825 CRC64;

Query Match 36.2%; Score 282; DB 2; Length 54;
Best Local Similarity 94.4%; Pred. No. 1e-15;
Matches 51; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERYESWTRETESTWLTYSADAPPSAAPPDGPAGGLHSG 54
Db 1 MCGGSRADAIERYESWTRETESTWLTYSADAPPSAAPPDGPAGGLHSG 54

RESULT 12
Q8VBS8 PRELIMINARY; PRT; 54 AA.
AC Q8VBS8;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE BAALC isoform 1-8.
DE Name=BaalC;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=12956/5VEVAC;
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin U.L., Leone G., Rush U.J., Plass C., Heimonen K.,
RA Mrozek K., Sill H., Knuttila S., Koltz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BALC, the human member of a novel mammalian neuroectoderm gene
RT lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AF371324; AAL50520.1; -.
DR MGD; MGI:1928704; BaalC.
DR Pfam; PF06989; BAALC_N; 1.
SQ SEQUENCE 54 AA; 5667 MW; FEE8C0BFBCEB829 CRC64;

Query Match 35.7%; Score 278; DB 2; Length 54;
Best Local Similarity 94.4%; Pred. No. 2.2e-15;
Matches 51; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERYESWTRETESTWLTYSADAPPSAAPPDGPAGGLHSG 54
Db 1 MCGGSRADAIERYESWTRETESTWLTYSADAPPSAAPPDGPAGGLHSG 54

RESULT 13
Q790N3 PRELIMINARY; PRT; 54 AA.
AC Q790N3;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DE BAALC isoform 1-8.
DE Name=BaalC;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

```

```

OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRague-Dawley;
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin U.L., Leone G., Rush U.J., Plass C., Heimonen K.,
RA Mrozek K., Sill H., Knuttila S., Koltz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BALC, the human member of a novel mammalian neuroectoderm gene
RT lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AF371325; AAL50521.1; -.
DR InterPro; IPR009728; BAALC_N.
DR Pfam; PF06989; BAALC_N; 1.
SQ SEQUENCE 54 AA; 5667 MW; FEE8C0BFBCEB829 CRC64;

Query Match 35.7%; Score 278; DB 2; Length 54;
Best Local Similarity 94.4%; Pred. No. 2.2e-15;
Matches 51; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERYESWTRETESTWLTYSADAPPSAAPPDGPAGGLHSG 54
Db 1 MCGGSRADAIERYESWTRETESTWLTYSADAPPSAAPPDGPAGGLHSG 54

RESULT 14
Q801V5 PRELIMINARY; PRT; 32 AA.
AC Q801V5;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DE SI:ZC215113.3 (Novel protein similar to human brain and acute
DE leukemia, cytoplasmic (BALC)) (Fragment).
DE Name=SI:ZC215113.3;
GN Brachydanio rerio (Zebrafish) (Danio rerio).
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Sehla H.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL807244; CAD87801.1; -.
DR InterPro; IPR009728; BAALC_N.
DR Pfam; PF06989; BAALC_N; 1.
FT NON_TER 32
SQ SEQUENCE 32 AA; 3663 MW; 916A5445D2637F7 CRC64;

Query Match 20.7%; Score 161; DB 2; Length 32;
Best Local Similarity 90.6%; Pred. No. 3.1e-06;
Matches 29; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERYESWTRETESTWLTYSADAPPSAAPPDGPAGGLHSG 32
Db 1 MCGGSRADAIERYESWTRETESTWLTYSADAPPSAAPPDGPAGGLHSG 32

RESULT 15
Q72H02 PRELIMINARY; PRT; 306 AA.
AC Q72H02;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DE Hypothetical protein.
DE OrderedAccession=TC1434;
OS Thermus thermophilus (strain HB27 / ATCC BAA-163 / DSM 7039).
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC Thermus.
OX NCBI_TaxID=262724;
RN [1]

```

RP SEQUENCE FROM N.A.
 RX PubMed-15064768;
 RA Henne A., Brueggemann H., Raasch C., Wlezer A., Hartech T.,
 RA Liebegang H., Johann A., Lienard T., Gohl O., Martinez-Arias R.,
 RA Jacobi C., Starkuviene V., Schlenczek S., Dencker S., Huber R.,
 RA Klensk H.-P., Kramer W., Merkl R., Gottschalk G., Fritz H.-J.,
 RT "The genome sequence of the extreme thermophile Thermus
 thermophilus.";
 RL Nat. Biotechnol. 22:547-553(2004).
 DR EMBL; AB017305; AAS61776.1; -.
 KW Complete proteome.
 SQ SEQUENCE 306 AA; 31927 MM; 6184FB203F4CE1FA CRC64;

Query Match 12.5%; Score 97.5; DB 2; Length 306;
 Best Local Similarity 30.6%; Pred. No. 5.6;
 Matches 26; Conservative 12; Mismatches 38; Indels 9; Gaps 2;

QY 36 PPSAAPPDSCPEAGLHSGMLEDLPSNGVPRSTAPGIPNPEKTNCE---TQCPNPQ 91
 DB 164 PPSAAPPDSCPEAGLHSGMLEDLPSNGVPRSTAPGIPNPEKTNCE---TQCPNPQ 91
 QY 92 SLSSGP-----LTQKONGLOTTEAK 111
 DB 224 ASRRGPGAGGTLREKEGALPTQAO 248

Search completed: April 12, 2005, 15:42:10
 Job time : 96.6667 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 15:15:13 ; Search time 103.667 Seconds

(without alignments)
540.967 Million cell updates/sec

Title: US-10-705-716A-8

Sequence: 1 MCGGSRADAIERYRYESWT.....VTENIQMDRSKRVTKNCIN 145

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :					A_Geneseqp_16Dec04:*
					1: geneseqp1980s:*
					2: geneseqp1990s:*
					3: geneseqp2000s:*
					4: geneseqp2001s:*
					5: geneseqp2002s:*
					6: geneseqp2003as:*
					7: geneseqp2003bs:*
					8: geneseqp2004s:*
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.					
SUMMARIES					
Result No.	Score	Query Match	Length	DB ID	Description
1	767	100.0	145	8	AD048479 Mouse PTH
2	746	97.3	145	8	AD048473 Rat PTH r
3	649	84.6	145	4	AAB95018 Human pro
4	649	84.6	145	5	AAO19498 HSI prote
5	649	84.6	145	6	ABR58646 Human can
6	649	84.6	145	7	AD031800 Human nov
7	649	84.6	145	7	AD046959 Brain and
8	649	84.6	145	8	AD048475 Human PTH
9	621.5	81.0	180	7	AD046961 Brain and
10	472.5	61.6	149	7	AD046963 Brain and
11	329.5	43.0	92	5	AD081902 Human dlo
12	294	38.3	54	8	AD048481 Mouse PTH
13	288	37.5	73	7	AD046962 Brain and
14	278	36.2	54	7	AD046960 Brain and
15	278	36.2	54	8	AD048477 Human PTH
16	272	35.5	80	7	AD046964 Brain and
17	140	18.3	25	7	AD046979 Brain and
18	99	12.9	18	7	AD046977 Brain and
19	89	11.6	16	8	AD048482 PTH respo
20	87.5	11.4	778	2	AA13456 Dufky rec
21	84.5	11.0	684	4	ABR6330 Drosophi
22	84	11.0	1433	5	ABP35624 Fungal ZB
23	83.5	10.9	742	5	ABR6153 Aspergill
24	82	10.7	574	5	ABG95028 Human tra
25	81.5	10.6	1100	4	AA64930 Shrimp wh

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

ALIGNMENTS

RESULT 1	AD048479	standard; protein; 145 AA.
ID	AD048479;	
AC	XX	
DT	12-AUG-2004	(first entry)
DE	Mouse PTH responsive gene protein.	
XX	PTH responsive gene; PARG; bone-forming; bone; bone density modulation; transgenic animal; osteopathic; gene therapy; osteoporosis; mouse; murine.	
KW	Mus sp.	
OS	WO2004044152-A2.	
PN	27-MAY-2004.	
PD	10-NOV-2003; 2003WO-US035655.	
XX	12-NOV-2002; 2002US-0425532P.	
PR	(AMHP) WYETH.	
PA	Robinson JA, Stojanovic-Suenic V, Bablj P, Murrills RJ;	
PI	WPI; 2004-420299/39.	
XX	N-PSDB; AD048478.	
DR	New nucleic acid fragment encoding a PARG polypeptide, useful in preparing a composition for diagnosing, treating or preventing bone related disorders, e.g., osteoporosis.	
PT	Claim 9; SEQ ID NO 8; 169pp; English.	
PS	The invention relates to a novel PTH responsive gene (PARG) fragment encoding a polypeptide. The invention further comprises: a chimeric construct comprising the isolated nucleic acid fragment operatively linked to suitable regulatory sequences; a host cell transformed with the chimeric construct; a vector comprising the nucleic acid fragment; obtaining a nucleic acid fragment encoding the polypeptide; a method for obtaining a polypeptide; detecting the presence of the nucleic acid fragment; an antibody that specifically binds to one or more epitopes of a PARG polypeptide; a composition for regulating bone-forming activity in a mammal comprising the nucleic acid fragment, polypeptide or antibody; an agent that alters the expression of PARG gene or polypeptide;	
CC	Adj70139 Human hea	
CC	Adi40859 Human kin	
CC	Abm83972 Human dia	
CC	Aam40296 Human pol	
CC	Aam40295 Human pol	
CC	Aam42081 Human pol	
CC	Aam42082 Human pol	
CC	Adg98982 Antagonis	
CC	Aau1506 Novel hum	
CC	Aaw01875 Neuronal	
CC	Aaw01884 Invertebr	
CC	Abb61076 Drosophi	
CC	Abj17935 Drosophi	
CC	Aau76758 Novel rec	
CC	Abb58571 Drosophi	
CC	Abj37991 D melano	
CC	Abg95028 Novel hum	
CC	Abg15607 Novel hum	
CC	Ab069719 Pseudom	
CC	Aay24750 Sarcophag	

CC determining whether an agent alters the expression of PAlGB mRNA;
CC screening agents for effectiveness in altering expression of the nucleic
CC acid fragment; screening for agents useful for treating bone related
CC disorders; evaluating the efficacy of a treatment of a bone related
CC disorder in a subject; identifying polypeptides capable of binding to
CC PAlGB; monitoring the effectiveness of treatment of a subject with a bone
CC related agent; a transgenic animal comprising the DNA; an animal model
CC for the study of bone density modulation comprising a first group of
CC animals composed of the transgenic animal and a second group of control
CC animals; studying bone mass determinants; studying the modulation of bone
CC mass; studying an effect of PAlGB on bone disorders; identifying an agent
CC for treating bone related disorders; identifying whether an agent which
CC has bone forming activity; and a stably transfected cell line comprising
CC two constructs, the first construct comprising a ligand binding domain
CC linked to a DNA binding domain which is linked to an activation domain
CC all of which expression is driven by a constitutive promoter, the second
CC construct comprising multiple copies of DNA binding elements linked to a
CC minimal promoter which is linked to PAlGB cDNA, where upon the addition
CC of chemical inducer, transcription of PAlGB gene is induced. The PAlGB
CC polynucleotide has osteopathic activity. The PTH responsive gene may be
CC used to treat disorders by gene therapy. The nucleic acid is useful in
CC preparing a composition for diagnosing, treating or preventing bone
CC related disorders, e.g., osteoporosis. This sequence represents a PTH
CC responsive gene protein of the invention.
XX

SQ Sequence 145 AA;

Query Match 100.0%; Score 767; DB 8; Length 145;
Best Local Similarity 100.0%; Pred. No. 2,8e-78;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERPRYESTWRETETSTWLTYYTDSALPSAAATDSGPEAGGLHAGVLEDEL 60
Db 1 MCGGSRADAIERPRYESTWRETETSTWLTYYTDSALPSAAATDSGPEAGGLHAGVLEDEL 60
QY 61 SSGVLRPAAPGGIANPEKKNCGTCCPNSSQSLTQKNGIWLATEAKRDAXKMSAR 120
Db 61 SSGVLRPAAPGGIANPEKKNCGTCCPNSSQSLTQKNGIWLATEAKRDAXKMSAR 120
QY 121 EVAINTENIRQMDRSKRVTKNCIN 145
Db 121 EVAINTENIRQMDRSKRVTKNCIN 145

RESULT 2

ADO48473 standard; protein; 145 AA.

ADO48473;

12-AUG-2004 (first entry)

Rat PTH responsive gene protein.

XX PTH responsive gene; PAlGB; bone-forming; bone; bone density modulation;
KW transgenic animal; osteopathic; gene therapy; osteoporosis; rat.
XX

OS Rattus sp.

WO2004044152-A2.

27-MAY-2004.

10-NOV-2003; 2003WO-US035655.

12-NOV-2002; 2002US-0425532P.

(AMHP) WYETH.

Robinson JA, Stojanovic-Suenilic V, Babi J, Murrills RJ,

WPI; 2004-420299/39.

N-PSDB; ADO48472.

XX New nucleic acid fragment encoding a PAlGB polypeptide, useful in
PT preparing a composition for diagnosing, treating or preventing bone
PT related disorders, e.g., osteoporosis.
XX
XX Claim 9; SEQ ID NO 2; 169pp; English.

XX The invention relates to a novel PTH responsive gene (PAlGB) fragment
CC encoding a polypeptide. The invention further comprises a chimeric
CC construct comprising the isolated nucleic acid fragment operatively
CC linked to suitable regulatory sequences, a host cell transformed with the
CC chimeric construct; a vector comprising the nucleic acid fragment;
CC obtaining a nucleic acid fragment encoding the polypeptide; a method for
CC fragmenting a polypeptide; detecting the presence of the nucleic acid
CC fragment; an antibody that specifically binds to one or more epitopes of
CC a PAlGB polypeptide; a composition for regulating bone-forming activity
CC in a mammal comprising the nucleic acid fragment, polypeptide or antibody
CC ; an agent that alters the expression of PAlGB gene or polypeptide;
CC determining whether an agent alters the expression of PAlGB mRNA;
CC screening agents for effectiveness in altering expression of the nucleic
CC acid fragment; screening for agents useful for treating bone related
CC disorders; evaluating the efficacy of a treatment of a bone related
CC disorder in a subject; identifying polypeptides capable of binding to
CC PAlGB; monitoring the effectiveness of treatment of a subject with a bone
CC related agent; a transgenic animal comprising the DNA; an animal model
CC for the study of bone density modulation comprising a first group of
CC animals composed of the transgenic animal and a second group of control
CC animals; studying bone mass determinants; studying the modulation of bone
CC mass; studying an effect of PAlGB on bone disorders; identifying an agent
CC for treating bone related disorders; identifying whether an agent which
CC has bone forming activity; and a stably transfected cell line comprising
CC two constructs, the first construct comprising a ligand binding domain
CC linked to a DNA binding domain which is linked to an activation domain
CC all of which expression is driven by a constitutive promoter, the second
CC construct comprising multiple copies of DNA binding elements linked to a
CC minimal promoter which is linked to PAlGB cDNA, where upon the addition
CC of chemical inducer, transcription of PAlGB gene is induced. The PAlGB
CC polynucleotide has osteopathic activity. The PTH responsive gene may be
CC used to treat disorders by gene therapy. The nucleic acid is useful in
CC preparing a composition for diagnosing, treating or preventing bone
CC related disorders, e.g., osteoporosis. This sequence represents a PTH
CC responsive gene protein of the invention.
XX

SQ Sequence 145 AA;

Query Match 97.3%; Score 746; DB 8; Length 145;
Best Local Similarity 97.2%; Pred. No. 6,6e-76;
Matches 141; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERPRYESTWRETETSTWLTYYTDSALPSAAATDSGPEAGGLHAGVLEDEL 60
Db 1 MCGGSRADAIERPRYESTWRETETSTWLTYYTDSALPSAAATDSGPEAGGLHAGVLEDEL 60
QY 61 SSGVLRPAAPGGIANPEKKNCGTCCPNSSQSLTQKNGIWLATEAKRDAXKMSAR 120
Db 61 SSGVLRPAAPGGIANPEKKNCGTCCPNSSQSLTQKNGIWLATEAKRDAXKMSAR 120
QY 121 EVAINTENIRQMDRSKRVTKNCIN 145
Db 121 EVAINTENIRQMDRSKRVTKNCIN 145

RESULT 3

AAB95018 standard; protein; 145 AA.

AAB95018;

26-JUN-2001 (first entry)

Human protein sequence SEQ ID NO:16726.

Human; primer; detection; diagnosis; antisense therapy; gene therapy.

```

XX OS Homo sapiens.
XX PN EPI074617-A2.
XX PD 07-FEB-2001.
XX PF 26-JUL-2000; 2000EP-00116126.
XX PR 29-JUL-1999; 99JP-00248036.
XX PR 27-AUG-1999; 99JP-00300253.
XX PR 11-JAN-2000; 2000JP-00118776.
XX PR 02-MAY-2000; 2000JP-00183767.
XX PR 09-JUN-2000; 2000JP-00241899.
XX PA (HELI-) HELIX RES INST.
XX PI Oca T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX DR WPI, 2001-318749/34.
XX PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
XX PT length cDNAs defined in the specification, and for the detection and/or
XX PT diagnosis of the abnormality of the proteins encoded by the full-length
XX PT cDNAs.
XX PS Claim 8; SEQ ID NO 16726; 2537pp + Sequence Listing; English.
XX CC The present invention describes primer sets for synthesizing 5602 full-
XX CC length cDNAs defined in the specification. Where a primer set comprises:
XX CC (a) an oligo-dt primer and an oligonucleotide complementary to the
XX CC complementary strand of a polynucleotide which comprises one of the 5602
XX CC nucleotide sequences defined in the specification, where the
XX CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX CC of an oligonucleotide comprising a sequence complementary to the
XX CC complementary strand of a polynucleotide which comprises a 5'-end
XX CC sequence and an oligonucleotide comprising a sequence complementary to a
XX CC polynucleotide which comprises a 3'-end sequence, where the
XX CC oligonucleotide comprises at least 15 nucleotides and the combination of
XX CC the 5'-end sequence and 3'-end sequence is selected from those defined in the
XX CC specification. The primer sets can be used in antisense therapy and in
XX CC gene therapy. The primers are useful for synthesizing polynucleotides,
XX CC particularly full-length cDNAs. The primers are also useful for the
XX CC detection and/or diagnosis of the abnormality of the proteins encoded by
XX CC the full-length cDNAs. The primers allow obtaining of the full-length
XX CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
XX CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
XX CC oligonucleotides, all of which are used in the exemplification of the
XX CC present invention
XX CC
XX SQ Sequence 145 AA;
XX
XX Query Match 84.6%; Score 649; DB 4; Length 145;
XX Best Local Similarity 83.4%; Pred. No. 6.3e-65;
XX Matches 121; Conservative 10; Mismatches 14; Indels 0; Gaps 0;
XX
XX QY 1 MCGGSRADAIIPRYESWTRETSTWLTFTYDSDALPSAAATDSGPEAGLHAGVLEDGL 60
XX DB 1 MCGGSRADAIIPRYESWTRETSTWLTFTYDSDALPSAAATDSGPEAGLHAGVLEDGL 60
XX
XX QY 61 SSGNGLRPAAPGAIANPEKKMCGTCCPNSONLSSGFLTKOKNGLWATEAKRDAXKMSAR 120
XX DB 61 PSNGVPRSTAPGAIIPNEKKTCNCETCCPNQSLSSGFLTKOKNGLQTEAKRDAXKMPAK 120
XX
XX QY 121 EVAINTENIRQMDRSKRVTKNCIN 145
XX DB 121 EVTINVTDSIQMDRSRRITKNCVN 145
XX
XX RESULT 4
XX AAO19498

```

```

XX ID AAO19498 standard; protein; 145 AA.
XX AC AAO19498;
XX XX
XX DT 20-DEC-2002 (first entry)
XX XX
XX DE HSI protein variant.
XX XX
XX KM HSI, variant; cancer; tumour; unigene cluster; cytostatic; metastasis;
XX KM EST; expressed sequence tag; colon cancer; stomach cancer; breast cancer;
XX KW HS169395; HS127144; HS2; HS132793; HS3.
XX XX
XX OS Unidentified.
XX PN DE10103694-A1.
XX PD 01-AUG-2002.
XX PF 26-JAN-2001; 2001DE-01003694.
XX PR 26-JAN-2001; 2001DE-01003694.
XX PA (DELB-) DELBRUECK CENT MOLEKULARE MEDIZIN MAX.
XX PI Bretz D, Kemmerer W;
XX DR WPI, 2002-644836/70.
XX DR N-PSDB; AAL50100.
XX PT Diagnosis and therapy of tumors, by determining expression rates of
XX PT specific expressed sequence tags of the unigene cluster, and subsequently
XX PT blocking their expression.
XX XX
XX CC Claim 10; Page 5; 10pp; German.
XX CC
XX CC The present invention relates to the use of expressed sequence tags
XX CC (ESTs), or variants, of the unigene cluster HS169395 (HS1), HS127144
XX CC (HS2) and/or HS132793 (HS3) for diagnosis and therapy of tumors, in
XX CC which their expression rates in tumour cells and/or lymph nodes are
XX CC determined. The EST sequences are useful as prognostic markers of
XX CC survival of cancer patients (high levels of EST-related mRNA are
XX CC associated with a poor prognosis, specifically correlated with
XX CC development of metastases); and for diagnosis and/or therapy of solid
XX CC tumours, particularly of colon, stomach and breast. The present sequence
XX CC is a variant of the HSI protein shown in the exemplification of the
XX CC invention
XX CC
XX SQ Sequence 145 AA;
XX
XX Query Match 84.6%; Score 649; DB 5; Length 145;
XX Best Local Similarity 83.4%; Pred. No. 6.3e-65;
XX Matches 121; Conservative 10; Mismatches 14; Indels 0; Gaps 0;
XX
XX QY 1 MCGGSRADAIIPRYESWTRETSTWLTFTYDSDALPSAAATDSGPEAGLHAGVLEDGL 60
XX DB 1 MCGGSRADAIIPRYESWTRETSTWLTFTYDSDALPSAAATDSGPEAGLHAGVLEDGL 60
XX
XX QY 61 SSGNGLRPAAPGAIANPEKKMCGTCCPNSONLSSGFLTKOKNGLWATEAKRDAXKMSAR 120
XX DB 61 PSNGVPRSTAPGAIIPNEKKTCNCETCCPNQSLSSGFLTKOKNGLQTEAKRDAXKMPAK 120
XX
XX QY 121 EVAINTENIRQMDRSKRVTKNCIN 145
XX DB 121 EVTINVTDSIQMDRSRRITKNCVN 145
XX
XX RESULT 5
XX ABR58646
XX ID ABR58646 standard; protein; 145 AA.
XX AC ABR58646;
XX XX
XX DT 09-JUL-2003 (first entry)

```

XX DE Human cancer related protein SEQ ID NO:303.
 XX XX
 KW Human; cancer; diagnosis; screening; modulator; leukemia; ischaemia;
 KW heart disease; atherosclerosis; endometriosis.
 XX XX
 OS Homo sapiens.
 XX PN WO2003025138-A2.
 XX XX
 PD 27-MAR-2003.
 XX XX
 PF 17-SEP-2002; 2002MO-US029560.
 XX XX
 PR 17-SEP-2001; 2001US-0323469P.
 XX PR 20-SEP-2001; 2001US-0323887P.
 PR 13-NOV-2001; 2001US-0350666P.
 PR 08-FEB-2002; 2002US-0355145P.
 PR 08-FEB-2002; 2002US-0355257P.
 PR 12-APR-2002; 2002US-0372246P.
 XX XX
 PA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX XX
 PI Afar D, Aziz N, Gish KC, Hevezi PA, Mack DH, Wilson KE,
 PI Zlotnick A;
 XX XX
 DR WPI: 2003-354600/33.
 DR N-PsDB; ACC72796.
 XX XX
 PT New genes that are up-regulated or down-regulated in cancers, useful as
 PT markers for diagnosing e.g. cancer, ischemia or heart diseases, or as
 PT therapeutic targets for screening drugs for treating these diseases.
 XX XX
 PS Claim 12; Page 753; 767pp; English.
 XX XX
 CC The present invention describes an isolated nucleic acid molecule, which
 CC comprises the sequence of any of the genes that are up-regulated or down-
 CC regulated in specific cancers (e.g. about 1031 genes up-regulated in
 CC acute lymphocytic leukemia). ACC72641 to ACC72860 represent cancer
 CC related gene nucleotide sequences which encode the proteins given in
 CC ABR8521 to ABR85709. Also described: (1) determining the presence or
 CC absence of a pathological cell in a patient; (2) an expression vector
 CC comprising a nucleic acid molecule described above; (3) a host cell
 CC comprising the vector; (4) an isolated polypeptide, which is encoded by
 CC the nucleic acid; (5) an antibody that specifically binds the polypeptide
 CC of (4); (6) specifically targeting a compound to a pathological cell in a
 CC patient by administering to the patient the antibody above; and (7) a
 CC drug screening assay. The nucleic acid is useful as diagnostic markers or
 CC therapeutic targets. In particular, the nucleic acid is useful for
 CC diagnosing a pathology, e.g. cancer (e.g. cancer of the bone marrow,
 CC bladder, brain, breast, cervix, colon/rectum, kidney, lung, ovary,
 CC pancreas, prostate, skin and uterus), wounds, ischemia, heart diseases,
 CC atherosclerosis and endometriosis. The nucleic acid is also useful in
 CC drug screening, particularly for identifying agents for treating these
 CC pathologies
 XX XX
 SQ Sequence 145 AA;
 XX XX
 Query Match 84.6%; Score 649; DB 6; Length 145;
 Best Local Similarity 83.4%; Fred. No. 6.3e-65;
 Matches 121; Conservative 10; Mismatches 14; Indels 0; Gaps 0;
 OY 1 MCCGSRADAIPEPRYYESWTRETESTWLTLYTSDALPSAAATDSGPBAGGLAGVLEGL 60
 DB 1 MCCGSRADAIPEPRYYESWTRETESTWLTLYTSDALPSAAATDSGPBAGGLAGVLEGL 60
 OY 61 SSNGVLRPAAPGGINPEKXWNGTQCPSNQLSSGPIITQXNGIMATEAKRDAKMSAR 120
 DB 61 PSNGVPRSTAPGGINPEKKTNCETQCPQSLSSGPIITQXNGIMATEAKRDAKMSAR 120
 OY 121 EVAIWTEINIKOMDSKRYTKNCIN 145
 DB 121 EVTIWVTDISIQOMDSRRRTKNCVN 145

RESULT 6
 ADCC31800
 ID ADCC31800 standard; protein; 145 AA.
 XX XX
 AC ADCC31800;
 XX XX
 DT 18-DEC-2003 (first entry)
 XX XX
 DE Human novel polypeptide sequence, SEQ ID NO:1882.
 XX XX
 KW Human; diagnostic; drug screening; forensics; gene mapping;
 KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
 KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
 KW ulcers; osteoporosis; autoimmune disease; cancer;
 KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
 KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vlnnary;
 KW antulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
 KW gene therapy; chromosome 8.
 XX XX
 OS Homo sapiens.
 XX XX
 PN WO2003029271-A2.
 XX XX
 PD 10-APR-2003.
 XX XX
 PF 24-SEP-2002; 2002MO-US030474.
 XX XX
 PR 24-SEP-2001; 2001US-0324631P.
 XX XX
 PA (HYSE-) HYSEQ INC.
 XX XX
 PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QH, Wang J, Wehrman T;
 PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
 PI Haley-Vicente D, Dymanc RT;
 XX XX
 DR WPI: 2003-371981/35.
 DR N-PsDB; ADCC30829.
 XX XX
 PT New polynucleotide and polypeptide useful for diagnosing, preventing or
 PT treating conditions such as neurodegenerative diseases, anemias, platelet
 PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 PT cancer.
 XX XX
 PS Claim 20; SEQ ID NO 1882; 1185pp; English.
 XX XX
 CC The invention relates to 971 novel human cDNA sequences (ADCC29919-
 CC ADCC30889) and the polypeptides they encode (ADCC30890-ADCC31860). The
 CC invention also relates to nucleic acid sequences over 99% identical with
 CC the novel human cDNAs. The invention additionally encompasses expression
 CC vectors and host cells comprising a nucleic acid of the invention; the
 CC recombinant production of a polypeptide of the invention; an antibody
 CC against a polypeptide of the invention; a method of detecting
 CC polynucleotides or polypeptides of the invention; and methods of
 CC identifying a compound which binds to a polypeptide of the invention. The
 CC invention further discloses methods of preventing, treating or
 CC ameliorating a medical condition; kits comprising polynucleotide probes
 CC and/or monoclonal antibodies for carrying out the methods of the
 CC invention; methods for the identification of compounds that modulate the
 CC expression or activity of the polynucleotide and/or polypeptide; and 767
 CC coding sequences corresponding to the cDNA sequences of the invention
 CC (ADCC31861-ADCC32627) and the polypeptides encoded by the cDNAs (ADCC32628
 CC -ADCC3394). The nucleic acids and polypeptides of the invention are
 CC useful in diagnostics, drug screening, forensics, gene mapping, in the
 CC identification of mutations responsible for genetic disorders or other
 CC traits, for assessing biodiversity, and in producing many other types of
 CC data and products dependent on DNA and amino acid sequences. They are
 CC also used for treating diseases such as Parkinson's disease, Alzheimer's
 CC disease and other neurodegenerative diseases, anaemia, platelet
 CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 CC cancer. The nucleic acids may also be used as hybridisation probes or
 CC primers, and in the recombinant production of a protein. The polypeptides

are also useful in generating antibodies, as molecular weight markers, and as food supplements. The present sequence represents a specifically claimed human polypeptide sequence of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at http://www.wipo.int/pub/publicated_pct_sequences.

Sequence 145 AA:

Query Match 84.6%; Score 649; DB 7; Length 145;
Best Local Similarity 83.4%; Pred. No. 6,3e-65;
Matches 121; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

QY 1 MCGGSGRADAIERYYSWTRETESTWLTYSDDLPSAAATDSGPEAGLHAGVLEDL 60
DB 1 MCGGSGRADAIERYYSWTRETESTWLTYSDDLPSAAATDSGPEAGLHAGVLEDL 60
QY 61 SSGVLRPAAPGGIANPEKKNCGTCCPNQSUNLSSGPLTKONGMLWATEAKRDKMSAR 120
DB 61 PSNGVPRSTAPGGIPNPEKKNCECTCCPNQSUNLSSGPLTKONGMLWATEAKRDKMSAR 120
QY 121 EVAINTENIRQMDRSKRVTKNCIN 145
DB 121 EVTINVTDSIQMDRSRRITKNCVN 145

RESULT 7

ADM46959 standard; protein; 145 AA.

ADM46959;

03-JUN-2004 (first entry)

Brain and Acute Leukemia, Cytoplasmic alternate protein #1.

acute myelogenous leukemia; gene expression; BAMLc;

chronic myelogenous leukemia; prostate cancer; Brain and Acute Leukemia;

Cytoplasmic; exon.

Homo sapiens.

Key Location/Qualifiers

FT MISC-difference 41 /note= "encoded by GCS"

MO2003040347-A2.

15-MAY-2003.

12-NOV-2002; 2002MO-US036375.

09-NOV-2001; 2001US-0348210P.

(OHIS) UNIV OHIO STATE RES FOUND.

Tanner SM, De La Chapell A;

WPI; 2003-441564/41.

N-PSDB; ADM46951.

Characterizing acute or chronic myelogenous leukemia, or prostate cancer in a patient comprising assaying for the overexpression of one or more BAMLc transcripts in cells obtained from the patient.

Disclosure; SEQ ID NO 17; 78bp; English.

The invention relates to a method of characterizing acute myelogenous leukemia (AML) in a patient by assaying for the overexpression of one or more BAMLc transcripts in cells obtained from the patient, where an overexpression indicates that the patient has an aggressive form of AML. The methods, kits and probes are useful for characterizing acute or chronic myelogenous leukemia, or prostate cancer. They are also useful

for detecting BAMLc overexpression. This sequence corresponds to a BAMLc (Brain and Acute Leukemia, Cytoplasmic) protein encoded by an alternative spliced RNA consisting of exons 1, 6 and 8.

Sequence 145 AA:

Query Match 84.6%; Score 649; DB 7; Length 145;
Best Local Similarity 83.4%; Pred. No. 6,3e-65;
Matches 121; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

QY 1 MCGGSGRADAIERYYSWTRETESTWLTYSDDLPSAAATDSGPEAGLHAGVLEDL 60
DB 1 MCGGSGRADAIERYYSWTRETESTWLTYSDDLPSAAATDSGPEAGLHAGVLEDL 60
QY 61 SSGVLRPAAPGGIANPEKKNCGTCCPNQSUNLSSGPLTKONGMLWATEAKRDKMSAR 120
DB 61 PSNGVPRSTAPGGIPNPEKKNCECTCCPNQSUNLSSGPLTKONGMLWATEAKRDKMSAR 120
QY 121 EVAINTENIRQMDRSKRVTKNCIN 145
DB 121 EVTINVTDSIQMDRSRRITKNCVN 145

RESULT 8

ADO48475 standard; protein; 145 AA.

ADO48475;

12-AUG-2004 (first entry)

Human PTH responsive gene protein.

PTH responsive gene; PARGB, bone-forming; bone, bone density modulation;

transgenic animal; osteopathic; gene therapy; osteoporosis; human.

Homo sapiens.

MO2004044152-A2.

27-MAY-2004.

10-NOV-2003; 2003MO-US035655.

12-NOV-2002; 2002US-0425532P.

(AMHP) WYETH.

Robinson JA, Stojanovic-Susulic V, Babij P, Murrill RJ;

WPI; 2004-420299/39.

N-PSDB; ADO48474.

New nucleic acid fragment encoding a PARGB polypeptide, useful in preparing a composition for diagnosing, treating or preventing bone related disorders, e.g., osteoporosis.

Claim 9; SEQ ID NO 4; 169bp; English.

The invention relates to a novel PTH responsive gene (PARGB) fragment encoding a polypeptide. The invention further comprises a chimeric construct comprising the isolated nucleic acid fragment operatively linked to suitable regulatory sequences; a host cell transformed with the chimeric construct; a vector comprising the nucleic acid fragment; obtaining a polypeptide; detecting the presence of the nucleic acid fragment; an antibody that specifically binds to one or more epitopes of a PARGB polypeptide; a composition for regulating bone-forming activity in a mammal comprising the nucleic acid fragment, polypeptide or antibody; an agent that alters the expression of PARGB gene or polypeptide; determining whether an agent alters the expression of PARGB mRNA; screening agents for effectiveness in altering expression of the nucleic acid fragment; screening for agents useful for treating bone related

CC disorders; evaluating the efficacy of a treatment of a bone related
 CC disorder in a subject; identifying polypeptides capable of binding to
 CC PAIGB; monitoring the effectiveness of treatment of a subject with a bone
 CC related agent; a transgenic animal comprising the DNA, an animal model
 CC for the study of bone density modulation comprising a first group of
 CC animals composed of the transgenic animal and a second group of control
 CC animals; studying bone mass determinants; studying the modulation of bone
 CC mass; studying an effect of PAIGB on bone disorders; identifying an agent
 CC for treating bone related disorders; identifying whether an agent which
 CC has bone forming activity; and a stably transfected cell line comprising
 CC two constructs, the first construct comprising a ligand binding domain
 CC linked to a DNA binding domain which is linked to an activation domain
 CC all of which expression is driven by a constitutive promoter, the second
 CC construct comprising multiple copies of DNA binding elements linked to a
 CC minimal promoter which is linked to PAIGB cDNA, where upon the addition
 CC of chemical inducer, transcription of PAIGB gene is induced. The PAIGB
 CC polynucleotide has osteopathic activity. The PTH responsive gene may be
 CC used to treat disorders by gene therapy. The nucleic acid is useful in
 CC preparing a composition for diagnosing, treating or preventing bone
 CC related disorders, e.g., osteoporosis. This sequence represents a PTH
 CC responsive gene protein of the invention.

CC Sequence 145 AA;

Query Match 84.6%; Score 649; DB 8; Length 145;
 Best Local Similarity 83.4%; Pred. No. 6.3e-65;
 Matches 121; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERPRYESWTRETSTWLTYYTSDALPSAAATDSGPEAGGLHAGVLEDEL 60
 DB 1 MCGGSRADAIERPRYESWTRETSTWLTYYTSDAPPSAAAPDSGPEAGGLHSGMLEDEL 60
 QY 61 SSGVLRPAAPGAIPEKKNMGCTOCPSNOMSSGGLTOKONGLWATEAKKDKAKMSAR 120
 DB 61 PSNGVPRSTAPGCIPEPEKKNMCTOCPSNOMSSGGLTOKONGLWATEAKKDKAKMPAK 120
 QY 121 EVAINTENIRQMDRSKRVTKNCIN 145
 DB 121 EVTINTDSIQQMDRSRRITTKNCVN 145

RESULT 9

ID ADM46961 standard; protein; 180 AA.

AC ADM46961;

DT 03-JUN-2004 (first entry)

DE Brain and Acute Leukemia, Cytoplasmic alternate protein #3.

KW acute myelogenous leukemia; gene expression; BAALC;
 KW chronic myelogenous leukemia; prostate cancer; Brain and Acute Leukemia;
 KW Cytoplasmic; exon.

OS Homo sapiens.

Key Location/Qualifiers

FT Misc-difference 41 /note= "encoded by GCS"

PN WO2003040347-A2.

PD 15-MAY-2003.

PF 12-NOV-2002; 2002MO-US036375.

PR 09-NOV-2001; 2001US-0348210P.

PA (OHIS) UNIV OHIO STATE RES FOUND.

PI Tanner SM, De La Chapell A;

XX

DR WPI; 2003-441564/41.

DR N-PSDB; ADM46953.

PT Characterizing acute or chronic myelogenous leukemia, or prostate cancer

PT in a patient comprises assaying for the overexpression of one or more

PT BAALC transcripts in cells obtained from the patient.

PS Disclosure; SEQ ID NO 19; 78bp; English.

XX The invention relates to a method of characterizing acute myelogenous
 CC leukemia (AML) in a patient by assaying for the overexpression of one or
 CC more BAALC transcripts in cells obtained from the patient, where an
 CC overexpression indicates that the patient has an aggressive form of AML.
 CC The methods, kits and probes are useful for characterizing acute or
 CC chronic myelogenous leukemia, or prostate cancer. They are also useful
 CC for detecting BAALC overexpression. This sequence corresponds to a BAALC
 CC (Brain and Acute Leukemia, Cytoplasmic) protein encoded by an alternative
 CC spliced RNA consisting of exons 1, 6 and 8.

CC Sequence 180 AA;

Query Match 81.0%; Score 621.5; DB 7; Length 180;
 Best Local Similarity 67.2%; Pred. No. 1.1e-61;
 Matches 121; Conservative 10; Mismatches 14; Indels 35; Gaps 1;

QY 1 MCGGSRADAIERPRYESWTRETSTWLTYYTSDALPSAAATDSGPEAGGLHAGVLEDEL 53
 DB 1 MCGGSRADAIERPRYESWTRETSTWLTYYTSDAPPSAAAPDSGPEAGGLHSGVLEAKS 60
 QY 54 -----GVLEGLSSNGVLRPAAPGAIPEKKNMGCT 85
 DB 61 KIKAPLTSDVSDGLPSASKMAPLAVFSGHMLEDGLPSNGVPRSTAPGCIPEPEKKNMCT 120
 QY 86 QCPNOMSSGGLTOKONGLWATEAKKDKAKMSAREVAINTENIRQMDRSKRVTKNCIN 145
 DB 121 QCPNOMSSGGLTOKONGLWATEAKKDKAKMPAKEVTINTDSIQQMDRSRRITTKNCVN 180

RESULT 10

ID ADM46963 standard; protein; 149 AA.

AC ADM46963;

DT 03-JUN-2004 (first entry)

DE Brain and Acute Leukemia, Cytoplasmic alternate protein #5.

KW acute myelogenous leukemia; gene expression; BAALC;
 KW chronic myelogenous leukemia; prostate cancer; Brain and Acute Leukemia;
 KW Cytoplasmic; exon.

OS Homo sapiens.

Key Location/Qualifiers

FT Misc-difference 41 /note= "encoded by GCS"

PN WO2003040347-A2.

PD 15-MAY-2003.

PF 12-NOV-2002; 2002MO-US036375.

PR 09-NOV-2001; 2001US-0348210P.

PA (OHIS) UNIV OHIO STATE RES FOUND.

PI Tanner SM, De La Chapell A;

WPI; 2003-441564/41.

DR N-PSDB; ADM46955.

XX

PT Characterizing acute or chronic myelogenous leukemia, or prostate cancer
 PT in a patient comprising assaying for the overexpression of one or more
 PT BAALC transcripts in cells obtained from the patient.
 XX
 XX Disclosure, SEQ ID NO 21, 78pp; English.
 XX
 CC The invention relates to a method of characterizing acute myelogenous
 CC leukemia (AML) in a patient by assaying for the overexpression of one or
 CC more BAALC transcripts in cells obtained from the patient, where an
 CC overexpression indicates that the patient has an aggressive form of AML.
 CC The methods, kits and probes are useful for characterizing acute or
 CC chronic myelogenous leukemia, or prostate cancer. They are also useful
 CC for detecting BAALC overexpression. This sequence corresponds to a BAALC
 CC (Brain and Acute Leukemia, Cytoplasmic) protein encoded by an alternative
 CC spliced RNA consisting of exons 1, 6 and 8.
 XX
 SQ Sequence 149 AA;
 Query Match 61.6%; Score 472.5; DB 7; Length 149;
 Best Local Similarity 65.3%; Pred. No. 6,4e-45;
 Matches 94; Conservative 3; Mismatches 12; Indels 35; Gaps 1;
 QY 1 MCGCGSRADAIEPRYSWTRETESTWLTYSDDALPSAATDSGPAGGILHA----- 53
 1 MCGCGSRADAIEPRYSWTRETESTWLTYSDDALPSAATDSGPAGGILHA----- 53
 DB 61 KIKAPDVSDEGLPSASKNAPLAVFSGHMLBDGLPSNGVPRSTAPGGINPEKKNCGT 85
 QY 86 QCPNSONLSSGGLTKOKNGLWATE 109
 121 QCPNSONLSSGGLTKOKNGLWATE 144
 DB
 RESULT 11
 ADQ81902
 ID ADQ81902 standard; protein; 92 AA.
 AC ADQ81902;
 XX
 DT 09-SEP-2004 (first entry)
 XX
 DE Human dioxigenase 10.12.
 XX
 KW Human; enzyme; dioxigenase 10.12; malignant tumour; inflammation;
 KW Immunological disease; haemopathy; HIV infection.
 XX
 OS Homo sapiens.
 XX
 PN CN1344798-A.
 XX
 PD 17-APR-2002.
 XX
 PF 29-SEP-2000; 2000CN-00125495.
 XX
 PR 29-SEP-2000; 2000CN-00125495.
 XX
 PA (SHAN-) SHANGHAI BIODOOR GENE DEV CO LTD.
 XX
 PI Mao Y, Xie Y;
 XX
 DR WPI, 2002-509506/55.
 XX
 DR N-PSDB; ADQ81901.
 XX
 PT New polypeptide human dioxigenase 10.12 and polynucleotides encoding this
 PT polypeptide, useful for treating various diseases, such as malignant
 PT tumours, inflammations, immunological diseases, hemopathy and HIV
 PT infection.
 XX
 XX Claim 1; SEQ ID NO 2; 33pp; Chinese.
 XX
 CC The present invention discloses a new kind of polypeptide, human

CC dioxigenase 10.12, polynucleotides encoding this polypeptide, a DNA
 CC recombination process to produce the polypeptide and antagonist against
 CC the polypeptide. The present invention also discloses the method of
 CC applying the polypeptide in treating various diseases, such as malignant
 CC tumours, inflammations, immunological diseases, haemopathy and HIV
 CC infection. The present sequence is the human dioxigenase 10.12.
 XX
 SQ Sequence 92 AA;
 Query Match 43.0%; Score 329.5; DB 5; Length 92;
 Best Local Similarity 63.4%; Pred. No. 5e-29;
 Matches 64; Conservative 8; Mismatches 10; Indels 19; Gaps 1;
 QY 45 GPBAGGLHGVLEDDGSSNGVLRPAPPGGIANPEKKNNGTCCPNSSONLSSGGLTKOKNG 104
 11 GPBAGGLHGVLEDDGSSNGVLRPAPPGGIANPEKKNNGTCCPNSSONLSSGGLTKOKNG 51
 DB 105 LMAEAKRDAKRMASREVAINTENTIRMDRSKRTVCNIN 145
 52 LQTEAKRDAKRMASREVAINTENTIRMDRSKRTVCNIN 92
 QY
 DB
 RESULT 12
 ADQ48481
 ID ADQ48481 standard; protein; 54 AA.
 AC ADQ48481;
 XX
 DT 12-AUG-2004 (first entry)
 XX
 DE Mouse PTH responsive gene protein exon 2 splice variant.
 XX
 KW PTH responsive gene; PAIGB; bone-forming; bone; bone density modulation;
 KW transgenic animal; osteopathic; gene therapy; osteoporosis; mouse;
 XX
 OS Mus sp.
 XX
 PN WO2004044152-A2.
 XX
 PD 27-MAY-2004.
 XX
 PF 10-NOV-2003; 2003WO-US035655.
 XX
 PR 12-NOV-2002; 2002US-0425532P.
 XX
 PA (AMHP) WYETH.
 XX
 PI Robinson JA, Stojanovic-Susulic V, Babij P, Murrills RJ;
 XX
 DR WPI, 2004-420299/39.
 XX
 DR N-PSDB; ADQ48480.
 XX
 PT New nucleic acid fragment encoding a PAIGB polypeptide, useful in
 PT preparing a composition for diagnosing, treating or preventing bone
 PT related disorders, e.g., osteoporosis.
 XX
 XX Claim 9; SEQ ID NO 10; 169pp; English.
 XX
 CC The invention relates to a novel PTH responsive gene (PAIGB) fragment
 CC encoding a polypeptide. The invention further comprises: a chimeric
 CC construct comprising the isolated nucleic acid fragment operatively
 CC linked to suitable regulatory sequences; a host cell transformed with the
 CC chimeric construct; a vector comprising the nucleic acid fragment;
 CC obtaining a nucleic acid fragment encoding the polypeptide; a method for
 CC obtaining a polypeptide; detecting the presence of the nucleic acid
 CC fragment; an antibody that specifically binds to one or more epitopes of
 CC a PAIGB polypeptide; a composition for regulating bone-forming activity
 CC in a mammal comprising the nucleic acid fragment, polypeptide or antibody
 CC; an agent that alters the expression of PAIGB gene or polypeptide;
 CC determining whether an agent alters the expression of PAIGB mRNA;
 CC screening agents for effectiveness in altering expression of the nucleic
 CC acid fragment; screening for agents useful for treating bone related

disorders; evaluating the efficacy of a treatment of a bone related disorder in a subject; identifying polypeptides capable of binding to PAIGB; monitoring the effectiveness of treatment of a subject with a bone related agent; a transgenic animal comprising the DNA; an animal model for the study of bone density modulation comprising a first group of animals composed of the transgenic animal and a second group of control animals; studying bone mass determinants; studying the modulation of bone mass; studying an effect of PAIGB on bone disorders; identifying an agent for treating bone related disorders; identifying whether an agent which has bone forming activity; and a stably transfected cell line comprising two constructs, the first construct comprising a ligand binding domain linked to a DNA binding domain which is linked to an activation domain all of which expression is driven by a constitutive promoter, the second construct comprising multiple copies of DNA binding elements linked to a minimal promoter which is linked to PAIGB cDNA, where upon the addition of chemical inducer, transcription of PAIGB gene is induced. The PAIGB polynucleotide has osteopathic activity. The PTH responsive gene may be used to treat disorders by gene therapy. The nucleic acid is useful in preparing a composition for diagnosing, treating or preventing bone related disorders, e.g., osteoporosis. This sequence represents a PTH responsive gene protein of the invention.

SO Sequence 54 AA;

Query Match 38.3%; Score 294; DB 8; Length 54;
Best Local Similarity 100.0%; Pred. No. 2.4e-25;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGGSRADAIIPRYESWTRTESTWLTYYTDSALPSAAATDGPAGGHLHAG 54
DB 1 MCGGSRADAIIPRYESWTRTESTWLTYYTDSALPSAAATDGPAGGHLHAG 54

RESULT 13

ADM46962
ID ADM46962 standard; protein; 73 AA.

AC ADM46962;

DT 03-JUN-2004 (first entry)

DE Brain and Acute Leukemia, Cytoplasmic alternate protein #4.

KW acute myelogenous leukemia; gene expression; BAALC;

KM chronic myelogenous leukemia; prostate cancer; Brain and Acute Leukemia;

KX Cytoplasmic; exon.

OS Homo sapiens.

Key Location/Qualifiers

FT Misc-difference 41

FT /note= "encoded by GCS"

PN WO2003040347-A2.

PD 15-MAY-2003.

PF 12-NOV-2002; 2002WO-US036375.

PR 09-NOV-2001; 2001US-0348210P.

PS (OHIS) UNIV OHIO STATE RES FOUND.

PI Tanner SM, De la Chapell A;

DR WPI; 2003-441564/41.

XX N-PSDB; ADM46954.

XX Characterizing acute or chronic myelogenous leukemia, or prostate cancer

XX in a patient comprises assaying for the overexpression of one or more

XX BAALC transcripts in cells obtained from the patient.

XX The invention relates to a method of characterizing acute myelogenous leukemia (AML) in a patient by assaying for the overexpression of one or more BAALC transcripts in cells obtained from the patient, where an overexpression indicates that the patient has an aggressive form of AML. The methods, kits and probes are useful for characterizing acute or chronic myelogenous leukemia, or prostate cancer. They are also useful for detecting BAALC overexpression. This sequence corresponds to a BAALC (Brain and Acute Leukemia, Cytoplasmic) protein encoded by an alternative spliced RNA consisting of exons 1, 6 and 8.

SO Sequence 73 AA;

Query Match 37.5%; Score 288; DB 7; Length 73;
Best Local Similarity 91.4%; Pred. No. 1.8e-24;
Matches 53; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MCGGSRADAIIPRYESWTRTESTWLTYYTDSALPSAAATDGPAGGHLHAG 58
DB 1 MCGGSRADAIIPRYESWTRTESTWLTYYTDSALPSAAATDGPAGGHLHAG 58

RESULT 14

ADM46960
ID ADM46960 standard; protein; 54 AA.

AC ADM46960;

DT 03-JUN-2004 (first entry)

DE Brain and Acute Leukemia, Cytoplasmic alternate protein #2.

KW acute myelogenous leukemia; gene expression; BAALC;

KM chronic myelogenous leukemia; prostate cancer; Brain and Acute Leukemia;

KX Cytoplasmic; exon.

OS Homo sapiens.

Key Location/Qualifiers

FT Misc-difference 41

FT /note= "encoded by GCS"

PN WO2003040347-A2.

PD 15-MAY-2003.

PF 12-NOV-2002; 2002WO-US036375.

PR 09-NOV-2001; 2001US-0348210P.

PS (OHIS) UNIV OHIO STATE RES FOUND.

PI Tanner SM, De la Chapell A;

DR WPI; 2003-441564/41.

XX N-PSDB; ADM46952.

XX Characterizing acute or chronic myelogenous leukemia, or prostate cancer

XX in a patient comprises assaying for the overexpression of one or more

XX BAALC transcripts in cells obtained from the patient.

XX Disclosure; SEQ ID NO 18; 78pp; English.

XX The invention relates to a method of characterizing acute myelogenous leukemia (AML) in a patient by assaying for the overexpression of one or more BAALC transcripts in cells obtained from the patient, where an overexpression indicates that the patient has an aggressive form of AML. The methods, kits and probes are useful for characterizing acute or chronic myelogenous leukemia, or prostate cancer. They are also useful for detecting BAALC overexpression. This sequence corresponds to a BAALC (Brain and Acute Leukemia, Cytoplasmic) protein encoded by an alternative spliced RNA consisting of exons 1, 6 and 8.

SQ Sequence 54 AA;
 Query Match 36.2%; Score 278; DB 7; Length 54;
 Best Local Similarity 94.4%; Pred. No. 1.6e-23;
 Matches 51; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MCGGSRADAIERPRYESTWTRETSTWLTYYTDSALPSAAATDGPAGGLHAG 54
 |||||
 DB 1 MCGGSRADAIERPRYESTWTRETSTWLTYYTDSALPSAAATDGPAGGLHAG 54
 |||||
 RESULT 15
 ADO48477
 ID ADO48477 standard; protein; 54 AA.
 AC ADO48477;
 DT 12-AUG-2004 (first entry)
 XX
 DE Human PTH responsive gene protein exon 2 splice variant.
 KW PTH responsive gene; PAlGB, bone-forming; bone; bone density modulation;
 KW transgenic animal; osteopathic; gene therapy; osteoporosis; human.
 XX
 OS Homo sapiens.
 XX
 PN MO200404152-A2.
 XX
 PD 27-MAY-2004.
 XX
 PP 10-NOV-2003; 2003MO-US035655.
 XX
 PR 12-NOV-2002; 2002US-0425532P.
 XX
 PA (AMHP) WYETH.
 XX
 PI Robinson JA, Stojanovic-Susulic V, Babić P, Murrills RJ;
 DR WPI, 2004-420299/39.
 DR N-PSDB; ADO48476.
 XX
 PT New nucleic acid fragment encoding a PAlGB polypeptide, useful in
 PT preparing a composition for diagnosis, treating or preventing bone
 PT related disorders, e.g., osteoporosis.
 PS
 PS Claim 9; SEQ ID NO 6; 169pp; English.
 CC The invention relates to a novel PTH responsive gene (PAlGB) fragment
 CC encoding a polypeptide. The invention further comprises: a chimeric
 CC construct comprising the isolated nucleic acid fragment operatively
 CC linked to suitable regulatory sequences; a host cell transformed with the
 CC chimeric construct; a vector comprising the nucleic acid fragment;
 CC obtaining a nucleic acid fragment encoding the polypeptide; a method for
 CC obtaining a polypeptide; detecting the presence of the nucleic acid
 CC fragment; an antibody that specifically binds to one or more epitopes of
 CC a PAlGB polypeptide; a composition for regulating bone-forming activity
 CC in a mammal comprising the nucleic acid fragment, polypeptide or antibody
 CC ; an agent that alters the expression of PAlGB gene or polypeptide;
 CC determining whether an agent alters the expression of PAlGB mRNA;
 CC screening agents for effectiveness in altering expression of the nucleic
 CC acid fragment; screening for agents useful for treating bone related
 CC disorders; evaluating the efficacy of a treatment of a bone related
 CC disorder in a subject; identifying polypeptides capable of binding to
 CC PAlGB; monitoring the effectiveness of treatment of a subject with a bone
 CC related agent; a transgenic animal comprising the DNA; an animal model
 CC for the study of bone density modulation comprising a first group of
 CC animals composed of the transgenic animal and a second group of control
 CC animals; studying bone mass determinants; studying the modulation of bone
 CC mass; studying an effect of PAlGB on bone disorders; identifying an agent
 CC for treating bone related disorders; identifying whether an agent which
 CC has bone forming activity; and a stably transfected cell line comprising
 CC two constructs, the first construct comprising a ligand binding domain
 CC linked to a DNA binding domain which is linked to an activation domain

CC all of which expression is driven by a constitutive promoter, the second
 CC construct comprising multiple copies of DNA binding elements linked to a
 CC minimal promoter which is linked to PAlGB cDNA, where upon the addition
 CC of chemical inducer, transcription of PAlGB gene is induced. The PAlGB
 CC polynucleotide has osteopathic activity. The PTH responsive gene may be
 CC used to treat disorders by gene therapy. The nucleic acid is useful in
 CC preparing a composition for diagnosis, treating or preventing bone
 CC related disorders, e.g., osteoporosis. This sequence represents a PTH
 CC responsive gene protein of the invention.

SQ Sequence 54 AA;
 Query Match 36.2%; Score 278; DB 8; Length 54;
 Best Local Similarity 94.4%; Pred. No. 1.6e-23;
 Matches 51; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MCGGSRADAIERPRYESTWTRETSTWLTYYTDSALPSAAATDGPAGGLHAG 54
 |||||
 DB 1 MCGGSRADAIERPRYESTWTRETSTWLTYYTDSALPSAAATDGPAGGLHAG 54
 |||||

Search completed: April 12, 2005, 15:37:19
 Job time : 104.667 secs

This Page Blank (uspto)


```

QY      89 NSQNLSS-GPLTQKONGLMATEAKRDAK 115
      412 :::::
Db      412 DDRSQSLGPHYDERATLGETHMEKDT 439

```

RESULT 3
US-08-374-077C-2
; Sequence 2, Application US/08374077C
Date: 11/11/2008

```

: GENERAL INFORMATION:
: APPLICANT: Hall, Linda M.
: APPLICANT: Ren, Dejian
: APPLICANT: Zheng, Wei
: APPLICANT: Dubaïd, Manuel Marcel Paul
: TITLE OF INVENTION: Genes Encoding an Invertebrate Alphaunit
: TITLE OF INVENTION: Calcium Channel Subunit
: NUMBER OF SEQUENCES: 57
: CORRESPONDENCE ADDRESS:

```

QY 3 CGSSRADAI~~EP~~RY--YESWTR~~ET~~ESTWLT~~YD~~SDALPSAATD~~SG~~PEAGGLHAGVLE~~D~~GL 60
||| ||| :| : :
D~~b~~ 207 CGGGGISA~~PP~~RLTPEEAW~~Q~~LQ~~P~~-----NSVT~~S~~AGS~~T~~INSS~~F~~SSG-----GR 249
||| ||| :| :| :| :| :| :|

DQ 61 SSSGVLPRAPFGGIANPEKKNNC-----GTQC-----PNS 90
 | :: ||
Db 250 DPNSSY--SAYGGDSSSNCSNCIDITGDNSTLGLGVGDVYSFIADCDNSEDDDGDPPNN 307

QY 91 QNLSGCP-TOKQGLMWATEAKRDAKRMSAREVAINTENIQMDRSKRVTKNC 143
| : | | | : | : | : | : | : | : |
Db 308 QDLSSQTLRTAAI VAAVAAAAKEQAQESLADCE-SFSDRRQDADEVDYRI IQDC 360

RESULT 4
US-08-895-590-2

GENERAL INFORMATION:
APPLICANT: Hall, Linda M.
APPLICANT: Ren, Dejian

? APPLICANT: Zheng, Wei
 ? APPLICANT: Zhubald, Mannel Marcel Paul
 ? TITLE OF INVENTION: Genes Encoding an Insect Calcium Channel
 ? NUMBER OF SEQUENCES: 101
 ? CORRESPONDENCE ADDRESS:

Query Match	10.5%	Score 80.5;	DB 3;	Length 2516;
Best Local Similarity	22.4%;	Pred. No.15;		
Matches 39;	Conservative 23;	Mismatches 59;	Indels 53;	Gaps 8;

QY 61 SSSGVLPAPPGSIANPEKKMNC-----GTC-----PNS 90
 : : : :
 Db 250 DNSSY--SAGVGDDSSSSNCNDITGDNSTLHGLGVGDVCSFIADCDNSEDDEGDPNN 307

Qy 91 QNLSSGPI-TQKNGLMWATEAKRDAKRMSAREVAIVNTENIRQMDSKRVTNKC 143
| : | | | | : : : : : : : :
Dd 308 QDLSSQTLRITAAIIVAATAAQAQEQGLADCE-SFSDRQDADDEDVRIIQDC 360

RESULT 5
US-09-539-879A-2

GENERAL INFORMATION:
APPLICANT: Hall, Linda M.

TITLE OF INVENTION: Genes Encoding an Invertebrate Alpha

```

Query Match 10.3%; Score 79; DB 4; Length 244;
Best Local Similarity 24.0%; Pred. No. 0.78;
Matches 42; Conservative 23; Mismatches 74; Indels 36; Gaps 6

Cy 2 CGCGSRDADIPRYEYSEWTRETESTWLT- - - - -YTDSDALPSAAATDSCPEAGGLHAGVY 56
Db 74 GCGASAPRPSAPRPSARNRSRSETPARSSPSPGTHSSRLPARAGGTSSSIGTASRLA 133
Cy 57 EDGSSNGVAPRPAAPGAIAPBEKKNC- - - - -GTCCPNSON- - - 92
Db 134 GAPIRLNG- - - - -PPSASSGSIAPSAIPHCARPOAHOROPRRPANSOSSPATRANSDONPA 191
Cy 93 - - - - -LSGGPLTQKONGIMATEAKRDAKRMASREVAIVNTENIRQMDRSKRYTKACIN 145
Db 192 CRLASGSRSTKSTSVASSGCVKPRRKRKRROGTSALRT- - - - -IRQART-GTSKPAIN 243

RESULT 8
US-09-121-964-1
: Sequence 1, Application US/09121964
: Patent No. 6124447
: GENERAL INFORMATION:
: APPLICANT: Nalcori, Shunji
: TITLE OF INVENTION: NOVEL ENZYME CATALYZING DEPHOSPHORYLATION
: FILE REFERENCE: 32290-144753
: CURRENT APPLICATION NUMBER: US/09/121,964
: CURRENT FILING DATE: 1998-07-24
: NUMBER OF SEQ ID NOS: 9
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1
: LENGTH: 724
: TYPE: PRT
: ORGANISM: Sarcophaga peregrina
US-09-121-964-1

Query Match 10.2%; Score 78; DB 3; Length 724;
Best Local Similarity 27.7%; Pred. No. 4.9;
Matches 31; Conservative 8; Mismatches 39; Indels 34; Gaps 4

```


TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-174

Query Match 9.8%; Score 75.5; DB 4; Length 305;
Best Local Similarity 27.1%; Pred. No. 2.7;
Matches 48; Conservative 16; Mismatches 72; Indels 41; Gaps 11;

QY 3 CGG-----SPADAIPEPRYSW---TRETSTWLYTSDALPSAAAT-DGCEAG 49
DB 20 CGGKSTENTDSRSSASTTVESTKATKSSSK-ATKSDAKSGTTTADSKATAS 78
QY 50 GLHAGVLEDLSSNGVLRPAAGIANPEKKM-CGTCQCN---SQNLSSGPILOKNGL 105
DB 79 STRKA-----ANNGSAEKQSPAKNANPDQANQVNLQNLMPGQGLPQAITTSQTNF 132
QY 106 W--ATEAKRDAKRMS---AREVAIVTE-NIRQMD-----RSKRVTKNCIN 145
DB 133 LTAATTSQADQNNFRVLYIAEKELIPVNDARVQULTFISFEKKTYSDAEAKVAVN 189

RESULT 12

US-09-408-647A-2
Sequence 2, Application US/09408647A
Patent No. 6399858

GENERAL INFORMATION:
APPLICANT: Kobayashi, Donald
TITLE OF INVENTION: Chitinase Gene from Stenotrophomonas
TITLE OF INVENTION: maltophilia
FILE REFERENCE: Rut-Cook 98-0090
CURRENT APPLICATION NUMBER: US/09/408,647A
CURRENT FILING DATE: 1999-08-26
PRIOR APPLICATION NUMBER: 60/098,036
PRIOR FILING DATE: 1998-08-27
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 700
TYPE: PRT
ORGANISM: Stenotrophomonas maltophilia
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)...(41)
NAME/KEY: DOMAIN
LOCATION: (196)...(290)
NAME/KEY: DOMAIN
LOCATION: (330)...(483)
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: Genbank No. 6399858 AF014950
DATABASE ENTRY DATE: 1997-09-23
US-09-408-647A-2

Query Match 9.8%; Score 75.5; DB 3; Length 700;
Best Local Similarity 22.5%; Pred. No. 9.1;
Matches 29; Conservative 15; Mismatches 58; Indels 27; Gaps 2;

QY 26 TWLYTSDALPSAAATDSGPEAGLHAGVLEDLSSNGVLRPAAGIANPEKKM--- 81
DB 159 TWANASAGSHTFRAVATDNNNAVTSATVSTYASNDTPPSVPGGLSPKXTATVN 218
QY 82 -----NCG-----TQCPNSQNLSSGPILOKNGIMATEAKRDARMS 118
DB 219 LVMSAATDNGSGGVAGDYVRNGSLVSGSPATQYDTGGLTASTAYTYVRARDNAGNAS 278
QY 119 AREVAIVNT 127
DB 279 ASGGSISVT 287

RESULT 13

US-09-328-352-8115

Sequence 8115, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 8115
LENGTH: 726
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-8115

Query Match 9.6%; Score 74; DB 4; Length 726;
Best Local Similarity 33.8%; Pred. No. 14;
Matches 26; Conservative 8; Mismatches 23; Indels 20; Gaps 4;

QY 2 GCCGSPADAIPEPRYSWTRTSTWLYTSDALPSA--AATDSC-----PEAGLIHAG 54
DB 181 GFGAGREDYWEPPDNDVWGDGKE--WLAHRNSEALAGSNLAATMGLIYVNP- 231
QY 55 VLEDGLSSNGVLRPAAP 71
DB 232 ---GPQASGDPRSAAP 244

RESULT 14

US-09-618-425-9
Sequence 9, Application US/09618425
Patent No. 6475744

GENERAL INFORMATION:
APPLICANT: Reppert, Steven M.
APPLICANT: Weaver, David R.
APPLICANT: Zylka, Mark
APPLICANT: Jin, Xiaowei
APPLICANT: Kume, Kazuhiko
APPLICANT: Sittam, Sathyanarayana
APPLICANT: Shearman, Lauren
TITLE OF INVENTION: METHODS FOR IDENTIFYING COMPOUNDS WHICH
TITLE OF INVENTION: MODULATE CIRCADIAN RHYTHM
FILE REFERENCE: 00786-428001
CURRENT APPLICATION NUMBER: US/09/618,425
CURRENT FILING DATE: 2000-07-18
PRIOR APPLICATION NUMBER: 60/203,005
PRIOR FILING DATE: 2000-05-10
PRIOR APPLICATION NUMBER: 60/145,363
PRIOR FILING DATE: 1999-07-22
NUMBER OF SEQ ID NOS: 13
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9
LENGTH: 1113
TYPE: PRT
ORGANISM: Mus musculus
US-09-618-425-9

Query Match 9.6%; Score 74; DB 4; Length 1113;
Best Local Similarity 25.0%; Pred. No. 27;
Matches 38; Conservative 22; Mismatches 70; Indels 22; Gaps 7;

QY 1 MCGGGS-----RADAIPEPRYSWTRTSTWLYTSDALPSAAATDSGPEAGL 51
DB 860 LGAAGSSELAFLVPANAPNPEPTTSGSORRVENNEAH--SHELPEFISRSSP-----L 913
QY 52 HAGVLEDLSSNGVLRPAAGIANPEKKMNC--GTQCNQNLSSGPI---TQKNGWL 106
DB 914 QNLQLEMPAPESADAVRRG-AGPDARHCVTGSGSRSRHCTSGELATATAHQESAA 972
QY 107 ATEAKRDARMSAREVAIVNTEN-IRQMDRSK 137
DB 973 ASGSSASSTYFSTVDYASEVSENRQRPPDRQR 1004

```
RESULT 15
US-09-538-092-723
; Sequence 723, Application US/09538092
; Patent No. 675314
; GENERAL INFORMATION:
; APPLICANT: Glot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 723
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number YOR007C
US-09-538-092-723

Query Match          9.6%; Score 73.5; DB 4; Length 346;
Best Local Similarity 29.4%; Pred. No. 5.6;
Matches 37; Conservative 11; Mismatches 51; Indels 27; Gaps 6;

QY      5 GSRADAIEPRVYESWTRERE-STWLTYYT-----DSDALPSAATDSG--PEAGGLHAGVL 56
DB      201 GDNATEAMKRDYIESAKKVEQSINLEKTVPEQSRDADVDASQGSAGSLPDLGSLGGGL 260
QY      57 EDGLSSNGVLRPAA-----PGGIANPEKKM-----CGTCCPNSQNLSSGPL 98
DB      261 -GGLMNNPQLMQAQQGMSNPGAMQNIQGMMDPSIRQMAEGFASGCGTPTNLSDLNNPDA 319
QY      99 TOKONG 104
DB      320 LRMVAG 325
```

Search completed: April 12, 2005, 15:43:30
Job time : 27.6667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 15:37:30 ; Search time 75.6667 Seconds
(without alignments)
636.206 Million cell updates/sec

Title: US-10-705-716a-8

Perfect score: 767

Sequence: 1 MCGGGRADAIERYESWT.....VTENIRQMDRSKVTKNKCN 145

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1418010 seqs, 33197259 residues

Total number of hits satisfying chosen parameters: 1418010

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: Published Applications AA:
2: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubppa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubppa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	649	84.6	145	US-10-293-239-17	Sequence 17, Appl
2	649	84.6	145	US-10-177-390-30	Sequence 10, Appl
3	621.5	81.0	180	US-10-293-239-19	Sequence 30, Appl
4	472.5	61.6	149	US-10-293-239-21	Sequence 21, Appl
5	288	37.5	73	US-10-293-239-20	Sequence 20, Appl
6	278	36.2	54	US-10-293-239-18	Sequence 18, Appl
7	272	35.5	80	US-10-293-239-22	Sequence 22, Appl
8	140	18.3	25	US-10-293-239-37	Sequence 37, Appl
9	99	12.9	18	US-10-293-239-35	Sequence 35, Appl
10	86	11.2	1597	US-10-437-963-180225	Sequence 180225,
11	84	11.0	147	US-10-767-701-180225	Sequence 48905, A
12	84	11.0	1433	US-09-801-368-60	Sequence 60, Appl
13	84	11.0	1433	US-10-149-310-112	Sequence 112, App

14	82.5	10.8	274	US-10-425-114-65588	Sequence 65588, A
15	82.5	10.8	609	US-10-437-963-144983	Sequence 144983,
16	82	10.7	455	US-10-424-599-193656	Sequence 193656,
17	82	10.7	1479	US-10-437-963-180217	Sequence 180217,
18	81	10.6	219	US-10-156-761-13447	Sequence 13447, A
19	80.5	10.5	600	US-10-408-765A-1945	Sequence 1945, Ap
20	80	10.4	528	US-10-437-963-173618	Sequence 173618,
21	80	10.4	1108	US-09-949-029-108	Sequence 108, App
22	79.5	10.4	996	US-10-437-963-114095	Sequence 114095,
23	78.5	10.3	876	US-09-801-368-44	Sequence 44, Appl
24	78.5	10.2	1274	US-10-437-963-180220	Sequence 180220,
25	78	10.2	1491	US-10-437-963-140197	Sequence 140197,
26	77.5	10.1	670	US-10-156-761-14107	Sequence 14107, A
27	77	10.0	268	US-10-369-493-8240	Sequence 8240, Ap
28	76.5	10.0	320	US-10-425-114-43318	Sequence 43318, Ap
29	76.5	10.0	426	US-10-369-493-3526	Sequence 4326, Ap
30	76	9.9	391	US-10-437-963-177907	Sequence 177907,
31	75.5	9.8	276	US-10-437-963-158275	Sequence 158275,
32	75.5	9.8	286	US-09-071-035-176	Sequence 176, App
33	75.5	9.8	286	US-10-206-576-176	Sequence 176, App
34	75.5	9.8	286	US-10-912-362-176	Sequence 176, App
35	75.5	9.8	305	US-09-071-035-174	Sequence 174, App
36	75.5	9.8	305	US-10-206-576-174	Sequence 174, App
37	75.5	9.8	305	US-10-912-362-174	Sequence 174, App
38	75.5	9.8	373	US-10-437-963-104701	Sequence 104701,
39	75	9.8	167	US-10-767-701-46805	Sequence 46805, A
40	75	9.8	604	US-10-437-963-163101	Sequence 163101,
41	74.5	9.7	286	US-10-437-963-163696	Sequence 163696,
42	74.5	9.7	580	US-10-282-122A-73164	Sequence 73164, A
43	74.5	9.7	601	US-10-282-122A-74859	Sequence 74859, A
44	74.5	9.7	601	US-10-282-122A-76267	Sequence 76267, A
45	74	9.6	332	US-10-437-963-168390	Sequence 168390,

ALIGNMENTS

US-10-293-239-17	RESULT 1
Sequence 17, Application US/10293239	
GENERAL INFORMATION:	
APPLICANT: Tanner, Stephen	
TITLE OF INVENTION: BAAC expression as a diagnostic marker for acute leukemia	
FILE REFERENCE: 22727/04101	
CURRENT APPLICATION NUMBER: US/10/293,239	
CURRENT FILING DATE: 2002-11-12	
PRIOR APPLICATION NUMBER: US 60/348,210	
PRIOR FILING DATE: 2001-11-09	
NUMBER OF SEQ ID NOS: 39	
SOFTWARE: PatentIn version 3.1	
SEQ ID NO 17	
LENGTH: 145	
TYPE: PRT	
ORGANISM: Homo sapiens	
US-10-293-239-17	
Query Match	84.6%; Score 649; DB 14; Length 145;
Best Local Similarity	83.4%; Pred. No. 8, 6e-06;
Matches 121; Conservative 10; Mismatches 14; Indels 0; Gaps 0;	
QY	1 MCGGGRADAIERYESWTRETSTWLTYSDDALPSAAADSGPAGGLAGVLEDL 60
DB	1 MCGGGRADAIERYESWTRETSTWLTYSDDALPSAAADSGPAGGLAGVLEDL 60
QY	61 SSGNGVRPAAPGGIAPPEKMKNGCTCPNSQNSASPELTKOKNGLWATEKPKAKMSAR 120
DB	61 PSNGVRPAAPGGIAPPEKMKNGCTCPNSQNSASPELTKOKNGLWATEKPKAKMSAR 120
QY	121 EVAINVTENIRQMDRSKVTKNKCN 145
DB	121 EVAINVTENIRQMDRSKVTKNKCN 145

```

RESULT 2
US-10-177-390-30
; Sequence 30, Application US/10177390
; Publication No. US20030143743A1
; GENERAL INFORMATION:
; APPLICANT: Schuler, Gerold
; APPLICANT: N.V. Antwerp Innovatiecentrum
; TITLE OF INVENTION: Improved transfection of Eucaryotic Cells with Linear
; TITLE OF INVENTION: Polynucleotides by Electroporation
; FILE REFERENCE: 021505wo/3H/ml
; CURRENT APPLICATION NUMBER: US/10/177.390
; CURRENT FILING DATE: 2002-06-20
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-390-30

```

Query Match	Similarity	Score	DB	Length
Best Local	83.4%	649	14	145
Matches	Conservative	10	Pre.Matches	14
			Indels	0
			Gaps	0

Qy	Db	Qy	Db	Qy	Db
1	1	61	61	121	121
MGCGGSRDAIEPRYESMTRETESTWTLYTTSQDALPSAAATADSPBEAGLHAGVLEDDL	MGCGGSRDAIEPRYESMTRETESTWTLYTTSQDALPSAAATADSPBEAGLHAGVLEDDL	SSNGVLRPAAPGIGIANPEKKONCGTQCENSONLSSGPIYLOKONGLMATEAKRDARMSAR	SSNGVLRPAAPGIGIANPEKKONCGTQCENSONLSSGPIYLOKONGLMATEAKRDARMSAR	EVAINVTENIRQMDRSKRYTKACIN	EVAINVTENIRQMDRSKRYTKACIN
1	1	61	61	121	121
MGCGGSRDAIEPRYESMTRETESTWTLYTTSQDALPSAAATADSPBEAGLHAGVLEDDL	MGCGGSRDAIEPRYESMTRETESTWTLYTTSQDALPSAAATADSPBEAGLHAGVLEDDL	PSNGVPRSTAGGINPEKNTNCETQCENPQSLGPIYLOKONGLOTTEAKRDARMPKX	PSNGVPRSTAGGINPEKNTNCETQCENPQSLGPIYLOKONGLOTTEAKRDARMPKX	EVAINVTENIRQMDRSKRYTKACIN	EVAINVTENIRQMDRSKRYTKACIN
1	1	61	61	121	121
MGCGGSRDAIEPRYESMTRETESTWTLYTTSQDALPSAAATADSPBEAGLHAGVLEDDL	MGCGGSRDAIEPRYESMTRETESTWTLYTTSQDALPSAAATADSPBEAGLHAGVLEDDL	SSNGVLRPAAPGIGIANPEKKONCGTQCENSONLSSGPIYLOKONGLMATEAKRDARMSAR	SSNGVLRPAAPGIGIANPEKKONCGTQCENSONLSSGPIYLOKONGLMATEAKRDARMSAR	EVAINVTENIRQMDRSKRYTKACIN	EVAINVTENIRQMDRSKRYTKACIN
1	1	61	61	121	121
MGCGGSRDAIEPRYESMTRETESTWTLYTTSQDALPSAAATADSPBEAGLHAGVLEDDL	MGCGGSRDAIEPRYESMTRETESTWTLYTTSQDALPSAAATADSPBEAGLHAGVLEDDL	PSNGVPRSTAGGINPEKNTNCETQCENPQSLGPIYLOKONGLOTTEAKRDARMPKX	PSNGVPRSTAGGINPEKNTNCETQCENPQSLGPIYLOKONGLOTTEAKRDARMPKX	EVAINVTENIRQMDRSKRYTKACIN	EVAINVTENIRQMDRSKRYTKACIN

```

RESULT 3
US-10-293-239-19
; Sequence 19, Application US/10293239
; Publication No. US20030119043A1
; GENERAL INFORMATION:
; APPLICANT: Tanner, Stephan
; APPLICANT: de la Chapelle, Albert
; TITLE OF INVENTION: BALC expression as a diagnostic marker for acute leukemia
; FILE REFERENCE: 22727/04101
; CURRENT APPLICATION NUMBER: US/10/293,239
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/348,210
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 180
; TYPE: PRY
; ORGANISM: Homo sapiens
; US-10-293-239-19

```

	Query Match	81.0%	Score 621.5,	DB 14,	Length 189;	
	Best Local Similarity	67.2%;	Pred. No. 8;6-e-57;			
	Matches 121;	Conservative 10;	Mismatches 14;	Indels 35;	Gaps 1	
OY	1 MGGCGSRADALEPRYESWTRSTETSLTWLYTDSDALPAAATDSEAGGJHA-----	53				
	: :					
Db	1 MGCGSSRADALAPRYESWTRBSTLWTLYTSDSPAAPAADSGEPAGEGLHSVLEAKS	60				
OY	54 -----GVTEEDGASNGVLRPAAPPGIANDPEKKNQCGT	85				
	: ::					
Db	61 KIRAPTDVSVDGLEFSASKMAPLAIVFSGHMLEDLPSNIGVPFRSTLPGGIPNPDEKKTNCET	120				
OY	86 QCPNSONLSGPTLTOKONGRLATEARKDAKRMSAREVAIVNTENIROMDRSKRVTKNCIN	145				

[illegible]

```

RESULT 4
US-10-293-239-21
; Sequence 21, Application US/10293239
; Publication NO. US20030119043h1
; GENERAL INFORMATION:
; APPLICANT: Tanner, Stephan
; TITLE OF INVENTION: BALC expression as a diagnostic marker for acute leukemia
; FILE REFERENCE: 22727/04101
; CURRENT APPLICATION NUMBER: US/10/293.239
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/348,210
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-293-239-21

```

	Query Match	61.6%	Score 472.5;	DB 14;	Length 149;		
	Best Local Similarity	65.3%;	Pred. No. 2.7e-41;				
	Matches	94;	Conservative	3;	Mismatches 12; Indels 35; Gaps 1		
Qy	1	MGCGGSRADAI	EPRYESWTR	ETESTTWLT	YDSALPSAAATDSCGPAGGLHA-----	53	
						:	
Dd	1	MCGGGRRDAI	EPRYESTRET	BSTWLT	YTIDSDAPPSAAA	PDSGPFAGLHSLVLEAKS 60	
Qy	54	-----	GVL	EDGIS	SNGVLRPAAPGGINPEKKONCGT 85		
						:	
Dd	61	KIKAPTDSV	DEGLFSA	RMAPLAVF	SHGMLEDGLP	SNGVPRTA	PGINPEKKTNCET 120
Qy	86	QCPSNSQL	SSGPI	TOKONGLMATE 109			
							:
Dd	121	QCNPQSLS	SGPI	TOKONGLOTTE 144			

```

RESULT 5
US-10-293-239-20
; Sequence 20, Application US/10293239
; Publication No. US20030119043A1
; GENERAL INFORMATION:
; APPLICANT: Tamer, Stephan
; APPLICANT: de la Chapelle, Albert
; TITLE OF INVENTION: BALC expression as a diagnostic marker for acute leukemia
; FILE REFERENCE: 22727/04101
; CURRENT APPLICATION NUMBER: US/10/293.239
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/348,210
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ. ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-239-20

```

Query Match	37.5%	Score 288	DB 14	Length 73						
Beet Local Similarity	91.4%	Pred. No. 2.2e-22								
Matches	53	Conservative	2	Indels 0; Gaps 0						
Qy	1	MGCGGSRADAI	BPYYESWTR	ETESTWLT	YTDSDAL	PSAAATD	SGPAAGL	HA	GVLED	58
bb	1	MGCGGSRADAI	BPYYESWTR	ETESTWLT	YTDSDAL	PSAAAPAD	SGPAAGL	HA	GVLED	58

RESULT 6

```

SOFTWARE: PatentIn version 3.1
SQ ID NO 37
LENGTH: 25
TYPE: PRT
ORGANISM: Homo sapiens
US-10-293-239-37

Query Match      18.3%; Score 140; DB 14; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.8e-07;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      7 RADAIEPRYESWTRTESTWLTYYT 31
|||||
Db      1 RADAIEPRYESWTRTESTWLTYYT 25

RESULT 9
US-10-293-239-35
Sequence 35, Application US/10293239
Publication No. US20030119043A1
GENERAL INFORMATION:
APPLICANT: Tanner, Stephan
APPLICANT: de la Chapelle, Albert
TITLE OF INVENTION: BACic expression as a diagnostic marker for acute leukemia
FILE REFERENCE: 22727/04101
CURRENT APPLICATION NUMBER: US/10/293,239
CURRENT FILING DATE: 2002-11-12
PRIOR APPLICATION NUMBER: US 60/3348,210
PRIOR FILING DATE: 2001-11-09
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PatentIn version 3.1
SEQ ID NO 35
LENGTH: 18
TYPE: PRT
ORGANISM: Homo sapiens
US-10-293-239-35

Query Match      12.9%; Score 99; DB 14; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.0023;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      9 DAIEPRYESWTRTEST 26
|||||
Db      1 DAIEPRYESWTRTEST 18

RESULT 10
US-10-437-963-180225
Sequence 180225, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovacic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Yong
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 180225
LENGTH: 1597
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_77610C.1.pep
US-10-437-963-180225

```

```

RESULT 12
US-09-801-368-60
: Sequence 60, Application US/09801368
: Patent No. US20020128250A1
: GENERAL INFORMATION:
: APPLICANT: Busby, Robert
: APPLICANT: Cali, Brian
: APPLICANT: Hecht, Peter
: APPLICANT: Holtzman, Doug
: APPLICANT: Madden, Kevin
: APPLICANT: Maxon, Mary
: APPLICANT: Milne, Todd
: APPLICANT: No. US20020128250A1man, Thea
: APPLICANT: Royer, John
: APPLICANT: Salama, Sofie
: APPLICANT: Sherman, Amir
: APPLICANT: Silva, Jeff
: APPLICANT: Summers, Eric
: TITLE OR INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
: FILE REFERENCE: 109272.147
: CURRENT APPLICATION NUMBER: US/09/801,368
: CURRENT FILING DATE: 2001-03-07
: PRIOR APPLICATION NUMBER: US 09/487,558
: PRIOR FILING DATE: 2000-01-19
: PRIOR APPLICATION NUMBER: US 60/160,587

```

```

RESULT 14
US-10-425-114-65588
; Sequence 65588, Application US/10425114
; Publication NO. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with

```

Search completed: April 12, 2005, 16:03:43
Job time : 76.6667 secs

/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 65588
/ LENGTH: 274
/ TYPE: PRT
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB4863-011-Fl12_Fl1.pep
US-10-425-114-65588

Query Match 10.8%; Score 82.5; DB 15; Length 274;
Best Local Similarity 34.4%; Pred. No. 4;
Matches 31; Conservative 8; Mismatches 32; Indels 19; Gaps 4;

QY 32 DSDALPSAATDGGPAGGLHAGVLDGLSNGVLRPAAPGCI--ANPEKKMNCCTGCP 88
DB 147 DSDDAAPAAAHHDDQPAVAVVAAGGLGGNSSSG-LPPAAGATAEAPESLSLSGLPAP 205
QY 89 -----NSQNLSSGPELTQKONG 104
DB 206 AAEPAEAAADDESRNSQGQAS-PLLEEGEG 234

RESULT 15
US-10-437-963-144983
/ Sequence 144983, Application US/10437963
/ Publication No. US2004012343A1
GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Wu, Wei
/ APPLICANT: Boukharov, Andrey A.
/ APPLICANT: Barbazuk, Brad
/ APPLICANT: Li, Ping
/ TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53221)B
/ CURRENT APPLICATION NUMBER: US/10/437,963
/ CURRENT FILING DATE: 2003-05-14
/ NUMBER OF SEQ ID NOS: 204966
/ SEQ ID NO 144983
/ LENGTH: 609
/ TYPE: PRT
/ ORGANISM: Oryza sativa
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (1)..(609)
/ OTHER INFORMATION: unsure at all Xaa locations
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT4530_45746C.1.pep
US-10-437-963-144983

Query Match 10.8%; Score 82.5; DB 16; Length 609;
Best Local Similarity 25.7%; Pred. No. 11;
Matches 38; Conservative 21; Mismatches 68; Indels 21; Gaps 5;

QY 8 ADNIEERYIESWTR-----ETESTWLTYYTSDALPSAATDGGPAGGL 51
DB 20 ADDEEEEBEATVRPVEVASESKDQEEVEEBEEDSDVDALDMLDAE-GPDGSGR 78
QY 52 HAGVLDGLSNGVLRPAAPGCI-ANPEKKMNCCTGCPNSQNLSSGPELTQKON--GLMAT 108
DB 79 PAAAPTAAGGAARPNHAGVLSRPFQPSNRTO-KLASHIRATPLEEWEGRMNVGMS 137
QY 109 EAKRDAKMSAREVAIVNTENIRQMDRS 136
DB 138 NSVTTAIRDSIRETAIGKTRNTEKADRA 165

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: April 12, 2005, 15:28:59 ; Search time 21.6667 Seconds
(without alignments)
643.912 Million cell updates/sec

Title: US-10-705-716A-8

Perfect score: 1 MCGGSGRADIERRYESWT.....VTENINQMDRSKRVTKNCIN 145

Sequence: 1 MCGGSGRADIERRYESWT.....VTENINQMDRSKRVTKNCIN 145

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	87.5	11.4	778	2 A35970	erythrocyte-binding
2	84	11.0	365	1 GNVSSC	genome polyprotein
3	84	11.0	1433	2 S54587	CAT8 protein - yea
4	79.5	10.4	571	2 T43456	hypothetical prote
5	79.5	10.4	1063	2 T03743	bifocal protein -
6	79.5	10.3	876	1 A57988	regulatory protein
7	78.5	10.2	600	2 C69899	conserved hypochet
8	78	10.2	269	2 T37073	hypothetical prote
9	77.5	10.1	848	2 G90128	elongation factor
10	77.5	10.1	2559	2 T09144	probable guanine n
11	77	10.0	781	2 T41551	hypothetical prote
12	76	9.9	516	2 JB0301	inulinase (EC 3.2.
13	76	9.9	645	2 S19156	serotonin receptor
14	74.5	9.7	601	2 AH0784	probable transpor
15	74	9.6	1113	2 T14260	period protein per
16	73.5	9.6	281	2 T02813	hypothetical prote
17	73.5	9.6	346	2 S61991	hypothetical prote
18	73.5	9.6	435	2 AG1028	preillin [imported
19	72.5	9.5	451	2 S13337	tubulin alpha-2 ch
20	72.5	9.5	85	2 H69775	hypothetical prote
21	72.5	9.5	1050	2 S54640	KCS1 protein - yea
22	72	9.4	424	2 T01383	GMPase-activating
23	72	9.4	4957	2 T03455	AR protein - huma
24	72	9.4	5262	2 T03454	AR protein - huma
25	71.5	9.3	3488	2 T34418	hypothetical prote
26	71	9.3	652	2 G82962	hypothetical prote
27	71	9.3	1840	2 T30250	GPI protein - mous
28	71	9.3	5327	2 T13564	microtubule-associ
29	70.5	9.2	251	2 C75521	cytochrome-related

30	70.5	9.2	351	2 A56387	helix-loop-helix t
31	70.5	9.2	381	2 A11846	aulfate-binding pr
32	70.5	9.2	502	2 S28080	gag polyprotein
33	70.5	9.2	512	2 E95902	probable sugar ABC
34	70.5	9.2	1343	2 AF0611	cell division prot
35	70.5	9.2	1996	2 F71405	probable TWV resis
36	70.5	9.2	3623	2 T08618	intrinsic factor-B
37	70	9.1	368	1 TVMSML	transforming prote
38	70	9.1	419	2 A36109	pregnancy-specific
39	70	9.1	475	2 C70966	hypothetical prote
40	70	9.1	500	2 F90133	t-complex protein
41	70	9.1	614	2 T06741	hypothetical prote
42	70	9.1	1573	2 T50113	3-dehydroquinatase
43	69.5	9.1	284	2 I51172	transcription fact
44	69.5	9.1	346	2 AF3161	phage-related inte
45	69.5	9.1	367	1 MHCH	Ig mu chain C regi

ALIGNMENTS

RESULT 1

A35970 erythrocyte-binding protein - Plasmodium knowlesi

C/Species: Plasmodium knowlesi

C/Date: 14-Dec-1990 #sequence_revision 14-Dec-1990 #text_change 09-Jul-2004

C/Accession: A35970

R/Adams, J.H.; Hudson, D.E.; Torii, M.; Ward, G.E.; Wellms, T.E.; Aikawa, M.; Miller, I.

Cell 63, 141-153, 1990

A/Title: The Duffy receptor family of plasmodium knowlesi is located within the microne-

A/Reference number: A35970, PMID:91004213, PMID:2170017

A/Accession: A35970

A/Status: Preliminary

A/Molecule type: DNA

A/Residues: 1-778 <ADA>

A/Cross-references: UNIPROT:P22545, GB:M68518, GB:M37513, NID:g160273, PID:g160274

C/Keywords: transmembrane protein

Query Match

Best Local Similarity 11.4%; Score 87.5; DB 2; Length 778;

Matches 28; Conservative 13; Mismatches 44; Indels 3; Gaps 3;

QY 31 TDSDALPSAATNSGPEAGLHAGVLEDDGSSGV-LRPAAPGGINPEKQNC-GTQCP 88

DB 352 TVSSDVPVSGKXGSGSTSHALAGENGVRHGTDTPEKDEKADPQKQDKQDT 411

QY 89 NSQNLSS-GPLTQKONGLMATEAKRDAK 115

DB 412 DDRSGSLSPHPTDERATLGETHMEKDT 439

RESULT 2

GNVSSC genome polyprotein - sugarcane mosaic virus (strain SC) (fragment)

N/Contents: carboxyl end of nuclear inclusion protein b; coat protein

C/Species: sugarcane mosaic virus, SCMV

C/Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 09-Jul-2004

R/Frenkel, M.J.; Jilka, J.M.; McKern, N.M.; Strike, P.M.; Clark Jr., J.M.; Shukla, D.D.

J. Gen. Virol. 72, 237-242, 1991

A/Title: Unexpected sequence diversity in the amino-terminal ends of the coat proteins

A/Reference number: PH0207, PMID:91132116, PMID:1993866

A/Accession: PH0207

A/Molecule type: genomic RNA

A/Residues: 1-365 <FRB>

A/Cross-references: UNIPROT:P25242; GB:D00948; NID:g2222123; PIDN:BAA00796.1; PID:g222212

C/Superfamily: tobacco etch virus genome polyprotein

C/Keywords: coat protein; inclusion protein b (fragment) #status predicted <IPB>

F/1-52/Product: nuclear inclusion protein b (fragment) #status predicted <IPB>

F/53-365/Product: coat protein #status predicted <COA>

Query Match

Best Local Similarity 11.0%; Score 84; DB 1; Length 365;

Matches 21; Conservative 13; Mismatches 44; Indels 3; Gaps 3;

	Matches	30; Conservative	23; Mismatches	80; Indels	6; Gaps	4;
QY	5	CSRADAIERRYESENTRTESTWTLYTSDALPSAAATDSCPEACGLHAGVL - EDGLSS	62			
Db	21	GIXEEIEE - KYFKEQFAKDLPGYLEIDYND - EVFHHQAGTVDAAGQGGGAGNAGTPPATGAAA	78			
QY	63	NCVLRLPAPAPGGIAPNEPKKNCNCTQCPNSQNLSSSGLTQKQGLWATTEAKRDAKRSAAEV	122			
Db	79	QGGAPRPATGAAAGPPTTQ - GSQLPQGGATGGGGAQTGAGGTGCVTGGQRDKVDAGATT	136			
QY	123	AINVTENIRQMDRSKRVTK	141			
Db	137	GKITVPKLLKAMSKKRRLPK	155			

RESULT 3

CAT8 protein yeast (*Saccharomyces cerevisiae*)
N:Alternate names: MSP8 protein; protein YMR021.06c; protein YMR280C
C:Species: *Saccharomyces cerevisiae*
C:Date: 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 16-Aug-2004
C:Accession: S54587; S48234; S61595; S49498
R:Pearson, D.; Bowman, S.
submitted to the EMBL Data Library, May 1995
A:Reference number: S54582
A:Accession: S54587
A:Molecule type: DNA
A:Residues: 1-1433 <PEA>
A:Cross-references: UNIPROT:P39113; EMBL:Z49704; NID:g825540; PIDN:CAAB9778.1; PID:g82555
A:Experimental source: strain AB972
R:Grezeltza, D.
submitted to the EMBL Data Library, March 1994
A:Reference number: S48234
A:Accession: S48234
A:Molecule type: DNA
A:Residues: 1-746, 'L', 748-1433 <GRZ>
A:Cross-references: EMBL:X78344; NID:g559523; PIDN:CAA55139.1; PID:g559524
R:Boles, E.; Hettmann, C.; Zimmermann, F.K.
submitted to the EMBL Data Library, December 1995
A:Reference number: S61594
A:Accession: S61595
A:Molecule type: DNA
A:Residues: 1-302, 'A', 304-767, 'A', 769-998, 1003-1007, 'S', 1009-1015, 'Q', 1017-1018, 'P', 1020-1021, 'D', 1023-1024, 'D', 1026-1027, 'D', 1029-1030, 'D', 1032-1033, 'D', 1035-1036, 'D', 1038-1039, 'D', 1041-1042, 'D', 1044-1045, 'D', 1047-1048, 'D', 1050-1051, 'D', 1053-1054, 'D', 1056-1057, 'D', 1059-1060, 'D', 1062-1063, 'D', 1065-1066, 'D', 1068-1069, 'D', 1071-1072, 'D', 1074-1075, 'D', 1077-1078, 'D', 1080-1081, 'D', 1083-1084, 'D', 1086-1087, 'D', 1089-1090, 'D', 1092-1093, 'D', 1095-1096, 'D', 1098-1099, 'D', 1101-1102, 'D', 1104-1105, 'D', 1107-1108, 'D', 1110-1111, 'D', 1113-1114, 'D', 1116-1117, 'D', 1119-1120, 'D', 1122-1123, 'D', 1125-1126, 'D', 1128-1129, 'D', 1131-1132, 'D', 1134-1135, 'D', 1137-1138, 'D', 1140-1141, 'D', 1143-1144, 'D', 1146-1147, 'D', 1149-1150, 'D', 1152-1153, 'D', 1155-1156, 'D', 1158-1159, 'D', 1161-1162, 'D', 1164-1165, 'D', 1167-1168, 'D', 1170-1171, 'D', 1173-1174, 'D', 1176-1177, 'D', 1179-1180, 'D', 1182-1183, 'D', 1185-1186, 'D', 1188-1189, 'D', 1191-1192, 'D', 1194-1195, 'D', 1197-1198, 'D', 1200-1201, 'D', 1203-1204, 'D', 1206-1207, 'D', 1209-1210, 'D', 1212-1213, 'D', 1215-1216, 'D', 1218-1219, 'D', 1221-1222, 'D', 1224-1225, 'D', 1227-1228, 'D', 1230-1231, 'D', 1233-1234, 'D', 1236-1237, 'D', 1239-1240, 'D', 1242-1243, 'D', 1245-1246, 'D', 1248-1249, 'D', 1251-1252, 'D', 1254-1255, 'D', 1257-1258, 'D', 1260-1261, 'D', 1263-1264, 'D', 1266-1267, 'D', 1269-1270, 'D', 1272-1273, 'D', 1275-1276, 'D', 1278-1279, 'D', 1281-1282, 'D', 1284-1285, 'D', 1287-1288, 'D', 1290-1291, 'D', 1293-1294, 'D', 1296-1297, 'D', 1299-1300, 'D', 1302-1303, 'D', 1305-1306, 'D', 1308-1309, 'D', 1311-1312, 'D', 1314-1315, 'D', 1317-1318, 'D', 1320-1321, 'D', 1323-1324, 'D', 1326-1327, 'D', 1329-1330, 'D', 1332-1333, 'D', 1335-1336, 'D', 1338-1339, 'D', 1341-1342, 'D', 1344-1345, 'D', 1347-1348, 'D', 1350-1351, 'D', 1353-1354, 'D', 1356-1357, 'D', 1359-1360, 'D', 1362-1363, 'D', 1365-1366, 'D', 1368-1369, 'D', 1371-1372, 'D', 1374-1375, 'D', 1377-1378, 'D', 1380-1381, 'D', 1383-1384, 'D', 1386-1387, 'D', 1389-1390, 'D', 1392-1393, 'D', 1395-1396, 'D', 1398-1399, 'D', 1401-1402, 'D', 1404-1405, 'D', 1407-1408, 'D', 1410-1411, 'D', 1413-1414, 'D', 1416-1417, 'D', 1419-1420, 'D', 1422-1423, 'D', 1425-1426, 'D', 1428-1429, 'D', 1431-1432, 'D', 1434-1435, 'D', 1437-1438, 'D', 1440-1441, 'D', 1443-1444, 'D', 1446-1447, 'D', 1449-1450, 'D', 1452-1453, 'D', 1455-1456, 'D', 1458-1459, 'D', 1461-1462, 'D', 1464-1465, 'D', 1467-1468, 'D', 1470-1471, 'D', 1473-1474, 'D', 1476-1477, 'D', 1479-1480, 'D', 1482-1483, 'D', 1485-1486, 'D', 1488-1489, 'D', 1491-1492, 'D', 1494-1495, 'D', 1497-1498, 'D', 1500-1501, 'D', 1503-1504, 'D', 1506-1507, 'D', 1509-1510, 'D', 1512-1513, 'D', 1515-1516, 'D', 1518-1519, 'D', 1521-1522, 'D', 1524-1525, 'D', 1527-1528, 'D', 1530-1531, 'D', 1533-1534, 'D', 1536-1537, 'D', 1539-1540, 'D', 1542-1543, 'D', 1545-1546, 'D', 1548-1549, 'D', 1551-1552, 'D', 1554-1555, 'D', 1557-1558, 'D', 1560-1561, 'D', 1563-1564, 'D', 1566-1567, 'D', 1569-1570, 'D', 1572-1573, 'D', 1575-1576, 'D', 1578-1579, 'D', 1581-1582, 'D', 1584-1585, 'D', 1587-1588, 'D', 1590-1591, 'D', 1593-1594, 'D', 1596-1597, 'D', 1599-1600, 'D', 1602-1603, 'D', 1605-1606, 'D', 1608-1609, 'D', 1611-1612, 'D', 1614-1615, 'D', 1617-1618, 'D', 1620-1621, 'D', 1623-1624, 'D', 1626-1627, 'D', 1629-1630, 'D', 1632-1633, 'D', 1635-1636, 'D', 1638-1639, 'D', 1641-1642, 'D', 1644-1645, 'D', 1647-1648, 'D', 1650-1651, 'D', 1653-1654, 'D', 1656-1657, 'D', 1659-1660, 'D', 1662-1663, 'D', 1665-1666, 'D', 1668-1669, 'D', 1671-1672, 'D', 1674-1675, 'D', 1677-1678, 'D', 1680-1681, 'D', 1683-1684, 'D', 1686-1687, 'D', 1689-1690, 'D', 1692-1693, 'D', 1695-1696, 'D', 1698-1699, 'D', 1701-1702, 'D', 1704-1705, 'D', 1707-1708, 'D', 1710-1711, 'D', 1713-1714, 'D', 1716-1717, 'D', 1719-1720, 'D', 1722-1723, 'D', 1725-1726, 'D', 1728-1729, 'D', 1731-1732, 'D', 1734-1735, 'D', 1737-1738, 'D', 1740-1741, 'D', 1743-1744, 'D', 1746-1747, 'D', 1749-1750, 'D', 1752-1753, 'D', 1755-1756, 'D', 1758-1759, 'D', 1761-1762, 'D', 176

	Query Match	Similarity	Score	DB 2	Length	1433
	Best Local	29.4%	Pred. No. 13			
	Matches	32	Conservative	15	Mismatches	50
					Indels	12
					Gaps	4
Qy	36	LP\$AATTSQGEAGGLHAGVLEDGLSSNGVLRPAPAGLAPNEKKNCGTQCPNQSS	95			
				:		:
Db	927	LP\$AATTSLLKPLFGSQSKNSLENRQRTPNVNR-----ENDEHYLLGNDSSNNNN	979			
				:		:
Qy	96	G-PLTQKQNGLMATEAKRDAKRMAREVAINVTENIR--QMDREKRYT	140			
				:		:
Db	980	GHSPEPTNTTNGKRLKYTEKDAK--NAKQGGISKGENAHNFQDYTKKNS	1027			
				:		:

RESULT 4

hypothetical protein DKFZp434I061.1 - human
C:Species: Homo sapiens (man)
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004

R|Accession: T43456
R|Poustka, A., Klein, M., Mewes, H.W., Gassenhuber, J., Wiemann, S
submitted to the Protein Sequence Database, December 1999
A|Reference number: Z22516
A|Accession: T43456
A|Status: preliminary
A|Molecule type: mRNA
A|Residues: 1-571 <AAA>
A|Cross-references: UNIPROT:O75175; EMBL:AL133647
A|Experimental source: adult testis, clone DKFZp34L061
A|Genetics:
A|Note: DKFZp34L061.1
A|Superfamily: hydroxyproline-rich glycoprotein

Query Match	10.4%	Score 79.5	DB 2	Length 571
Best Local Similarity	30.9%	Pred. No. 12		
Matches	29	Conservative	6	Mismatches 34; Indels 25; Gaps 3

QY	31	TDSDALPSAAATDSCGPEAGLH-----	-AGVLEDGLSNGVLEPPAA-	71
		:	:	
Db	111	TDSSVSGSPAKNGSKPVHSHNPQSPAVPPVPSGPPPAASALSTPGANGVAPAAAPPS		170
QY	72	--GGIANPEKKMNCGTCCPNSSONS----	-SGPLT	99
		: :		
Db	171	ALGPKASPAPBSHNSGTPAPYTAQAVAPAPPSCEPT		204

RESULT 5

b|focal protein- fruit fly (*Drosophila melanogaster*)
c|Species: *Drosophila melanogaster*
C|Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #ext_change 09-Jul-2004
C|Accession: T03743
R|Bahrl, S.M.; Yang, X.Y.; Chia, W.
Mol. Cell. Biol. 17, 5521-5529, 1997
A|Title: The *Drosophila* bifocal gene encodes a novel protein which colocalizes with actin
A|Reference number: Z15048; MOID:97415628; PMID:9271427
A|Accession: T03743
A|Status: preliminary; translated from GB/EMBL/DDbJ
A|Molecule type: mRNA
A|Residues: 1-1063 <BAH>
A|Cross-references: UNIPROT:Q16125; EMBL:AF011791; NID:g338667; PIDN:AAB6991.1; PID:g2
C|Genetics:
A|Cross-references: FlyBase:Fggn0014133
A|Note: bifocal

	Query Match	10.4%	Score 79.5;	DB 2;	Length 1063;
	Best Local Similarity	20.7%;	Pred. No. 25;		
	Matches	39;	Conservative	24;	Mismatches 58; Indels 67; Gaps 6;
QY	5	GSRADALEPRYSBWT-----ETESTWLTYYTSDAL-----PSAAT-	42		
		:: :::		::	
Db	83	GATADFTEPATISTSQKRWIGSEEEKSEKSSINTNSDSTGCHHSVAVSLSPDAATT	142		
		:: :::		::	
QY	43	-----DSGEAGLHNVGLDEDLSNGVLRLPAAG	72		
		:: :::		::	
Db	143	NVTYTPLEPKRSSLLNTRSGEREMRYILSBGSRDELBSGEOPAGVANSRCEVEYG	202		
		:: :::		::	
QY	73	GIANPEKKMNCGTCCPNSONLSGGPLTOXONGIMATEAKDKAMSAEVALINVTENIRQ	132		
		:: :::		::	
Db	203	TISGPSASNA--QNPNPHLK---TKCPQGQVAAKGKSACE-----TIVDNKS	247		
		:: :::		::	
QY	133	MDRSKRVT	140		
		:: :::		::	
Db	248	CSKTSTIS	255		
		:: :::		::	

RESULT 6

regulatory protein area - *Emericella nidulans*
C/Species: *Emericella nidulans*, *Aspergillus nidulans*
C/Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
C/Accession: A57988; S10017, S70168; S72883

R/Kudla, B.; Caddick, M.X.; Langdon, T.; Martinez-Rossi, N.M.; Bennett, C.F.; Sibley, S.
EMBO J. 9, 1355-1364, 1990
A/Title: The regulatory gene *areA* mediating nitrogen metabolite repression in *Aspergillus*
A/Reference number: S10017; MUID:90228331; PMID:1970293
A/Accession: A57988
A/Status: translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-876 <KUD>
A/Cross-references: UNIPROT:P17429; EMBL:X52491; NID:g1019911; PIDN:CAA36731.1; PID:g115
A/Note: this sequence represents reinterpretation to include two exons
A/Accession: S10017
A/Molecule type: DNA
A/Residues: 158-876 <KUD>
A/Cross-references: EMBL:X52491
A/Note: this sequence represents the authors' original translation
R/Langdon, T.; Sheerin, A.; Ravagnani, A.; Gielkens, M.; Caddick, M.X.; Arst Jr., H.N.
Mol. Microbiol. 17, 877-888, 1995
A/Title: Mutational analysis reveals dispensability of the N-terminal region of the *AspE*
A/Reference number: S70167; MUID:96123430; PMID:8596437
A/Accession: S70168
A/Status: nucleic acid sequence not shown
A/Molecule type: DNA
A/Residues: 1-791, 'A', '793', 'T', '795', 'SPGTNS', '802-876 <LAN>
A/Cross-references: EMBL:X52491
R/Caddick, M.X.
submitted to the EMBL Data Library, October 1995
A/Reference number: S72883
A/Accession: S72883
A/Molecule type: DNA
A/Residues: 1-876 <CAD>
A/Cross-references: EMBL:X52491; NID:g1019911; PIDN:CAA36731.1; PID:g1154625
C/Genetics:
A/Gene: *areA*
A/Intons: 147/2
C/Function:
A/Description: mediates nitrogen metabolite repression
C/Superfamily: nitrogen regulatory protein ntr-2; GATA-type zinc finger homology
C/Keywords: DNA binding; transcription regulation; zinc finger
P/670-723/Domain: GATA-type zinc finger homology <GZF>
P/673-697/Region: zinc finger GATA motif

Query Match 10.3%; Score 79; DB 1; Length 876;
Best Local Similarity 28.4%; Pred. No. 22;
Matches 31; Conservative 10; Mismatches 28; Indels 40; Gaps 5;

QY 57 EDGLSSNGVLRLPAPGSIANPEKK-----MNCGTQC-----PNSQNLSSG----- 96
DB 645 ESGINSAPRPARSPGCTKNGEQNGPTCTNCFQTTPLWRNPDEGQPLCNACGLFLKHL 704
QY 97 -----PLTKONGLMATKADAKMSAREVAIVNTENIRQMDSKRYTK 141
DB 705 GVVRRPLSLK-----TDVIKKRNNSANSLAVG-----SSRVSX 737

RESULT 7
conserved hypothetical protein yobL - *Bacillus subtilis*
C/Species: *Bacillus subtilis*
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C/Accession: C69899
R/Kunert, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertet
C.; Bron, S.; Brouillet, S.; Brunech, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerston, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A/Authors: Foulger, C.D.; Fritz, C.; Hilbert, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
lech, J.; Harwood, C.R.; Hentau, A.; Jilbert, H.; Holasappel, S.; Hosono, S.; Hullo, M.F.
Koetler, P.; Koningerstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidis, A.; Lardinois
A.; Luthers, L.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Nauee
Y., M.; Ogasawa, K.; Ogawa, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
R.; Rieger, M.; Rivolta, C.; Roche, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scallion
A.; Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serot
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpestre, P.; Tognoni, A.; Toasato, V.; Uchiyama
T.; Wintere, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.

A/Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A/Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*
A/Reference number: A69580; MUID:98044033; PMID:9384377
A/Accession: C69899
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-600 <KUN>
A/Cross-references: UNIPROT:Q34330; GB:Z99114; GB:AL009126; NID:g2634230; PIDN:CAB13792
A/Experimental source: strain 168
C/Genetics:
A/Gene: *yobL*

Query Match 10.2%; Score 78.5; DB 2; Length 600;
Best Local Similarity 23.7%; Pred. No. 16;
Matches 40; Conservative 22; Mismatches 56; Indels 51; Gaps 7;

QY 9 DAIEPRYYSWTR-----TESTWLYTSDALPSAAATSGPAGIAGVLDGLS 61
DB 342 DAISAAIEESYQDMVNGDAYSRSRWY-----AIGSVAVAVGTG-----AGAINKADA 393
QY 62 SNGVLRPAPAGIA-----NPEKKNCGTQC-----NSQNLSSGPT----- 99
DB 394 AGKVINKASQAGKKIKDKVLPDLLPYNPKYKALADNPYNVDSQNLKNEILLTNAKIP 453
QY 100 -----QKONGLMATEAKRDKMSAREVAIVNTENIRQMDSKRY 139
DB 454 DGTREKFTGQKSPPLNKEKYD-----AYEIGKVKAKGVADVSRV 497

RESULT 8
T37073
hypothetical protein SCJ30.06c - *Streptomyces coelicolor*
C/Species: *Streptomyces coelicolor*
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C/Accession: T37073
R/Sanders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.
submitted to the EMBL Data Library, August 1999
A/Reference number: Z21621
A/Accession: T37073
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-269 <SAN>
A/Cross-references: UNIPROT:Q9S1Y3; EMBL:AL109973; PIDN:CAB53301.1; GSPDB:GN00070; SCOB
A/Experimental source: strain A312)
C/Genetics:
A/Gene: SCOEDB:SCJ30.06c

Query Match 10.2%; Score 78; DB 2; Length 269;
Best Local Similarity 25.8%; Pred. No. 6.9;
Matches 25; Conservative 10; Mismatches 38; Indels 24; Gaps 4;

QY 11 IEBRYYESWT-----RETESWLYTSDALPSAAATSGPAGIAGVLDGL 60
DB 176 LPEPTHEQWVNVFHCRRPRRVLRSGGLTFSPVTVPSRGCCSGSRTG----- 225
QY 61 SSNGVLRPAPAGSIANPEKKMNG--TQCPNSQNLSS 95
DB 226 GGGVVTGAGGIV-RGEACRCGRAAHAPYNAKMS 260

RESULT 9
G90128
elongation factor EF-2 [imported] - *Gulliardia cheta* nucleomorph
C/Species: nucleomorph *Gulliardia cheta*
A/Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C/Accession: G90128
R/Douglass, S.; Zanner, S.; Fruhmholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Re
Nature 410, 1091-1096, 2001
A/Title: The highly reduced genome of an enslaved algal nucleus.
A/Reference number: A99082; MUID:11323671; PMID:11323671
A/Accession: G90128
A/Status: preliminary

A/Title: A family of Drosophila serotonin receptors with distinct intracellular signals
 A/Reference number: S19155; PMID:92155185; PMID:1310937
 A/Accession: S19156
 A/Molecule type: mRNA
 A/Residues: 1-645 <SAU>
 A/Cross-references: UNIPROT:P28286; EMBL:Z11490; NID:g7506; PIDN:CAA77571.1; PID:g7507
 C/Genetics:
 A/Gene: FlyBase:5-HT1B
 A/Cross-references: FlyBase:Fgn0004572
 C/Superfamily: Octopamine receptor type I
 C/Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
 F/124-145/Domain: transmembrane #status predicted <TM1>
 F/156-177/Domain: transmembrane #status predicted <TM2>
 F/193-214/Domain: transmembrane #status predicted <TM3>
 F/224-256/Domain: transmembrane #status predicted <TM4>
 F/284-305/Domain: transmembrane #status predicted <TM5>
 F/564-587/Domain: transmembrane #status predicted <TM6>
 F/597-619/Domain: transmembrane #status predicted <TM7>

Query Match 9.9%; Score 76; DB 2; Length 645;
 Best Local Similarity 31.0%; Pred. No. 30;
 Matches 31; Conservative 12; Mismatches 31; Indels 26; Gaps 4;

QY 29 TYTDSALPSAATDSQPEAGLHAGVLEDGLSSNGVLRPAAPG-----IANPEKKNCG 84
 DB 489 TTPPEKALSGAGTAVAGAVAGSGSGSGEGAGTEGKAGVGLGVLASTIANPHOKL--- 545
 QY 85 TQCPNSQNTSSGPLTQKNGLMATEAKRDARMSAREVAI 124
 DB 546 -----AKRRQL--LEAKRE--RKAQQLAI 566

RESULT 14

AH0784
 Probable transport system periplasmic binding protein STY2452 [imported] - Salmonella enterica
 C/Species: Salmonella enterica subsp. enterica serovar Typh
 A/Note: this species has also been called Salmonella typhi
 C/Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C/Accession: AH0784
 R/Parikh, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Conerton, P.; Cronin, A.; Davies, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Garra, P.
 Nature 413, 848-852, 2001
 A/Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A/Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov. A/Reference number: AB0502; PMID:21534947; PMID:11677608
 A/Accession: AH0784
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-601 <PAR>
 A/Cross-references: GB:AL513382; PIDN:CAM02598.1; PID:G16503454; GSPDB:GN00176
 C/Genetics:
 A/Gene: STY2452
 C/Superfamily: periplasmic oligopeptide-binding protein

Query Match 9.7%; Score 74.5; DB 2; Length 601;
 Best Local Similarity 22.8%; Pred. No. 38;
 Matches 33; Conservative 20; Mismatches 53; Indels 39; Gaps 7;

QY 15 YVESWTR-----ETESTWLTYTDSDA-----LPSAATD--SGPEAGLHAGVL 56
 DB 342 FYNWSRTNYSFONTEYAANNYPADDELVLAPMKDLPPEVFTQIYQPVNSG----- 395
 QY 57 EDGLSSNGVLRPA-----PGIANPEKKNCGTQCPNSQNTSSGPLTQKNGLMATEAKR 112
 DB 396 -DGYDRENILKADALLTQAGVINGQQRVNSVTKKPLTFELL--LPASSNQWVLPFOH 451
 QY 113 DAKMSAREVAIVNTENIRQMDRSK 137
 DB 452 NLQR-----LGITWTRQVDSQ 469

RESULT 15

T14260
 Period protein Per3 - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
 C/Accession: T14260
 R/Zylka, M.J.; Shearman, L.P.; Weaver, D.R.; Reppert, S.M.
 Neuron 20, 1103-1110, 1998
 A/Title: Three period homologs in mammals: differential light responses in the suprachiasmatic nucleus
 A/Reference number: Z17943; PMID:98318231; PMID:9655499
 A/Accession: T14260
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-1113 <ZYL>
 A/Cross-references: UNIPROT:O70361; EMBL:AF050182; NID:g3136149; PID:g3136150; PIDN:AMC4
 C/Genetics:
 A/Gene: Per3
 C/Keywords: circadian rhythm

Query Match 9.6%; Score 74; DB 2; Length 1113;
 Best Local Similarity 25.0%; Pred. No. 87;
 Matches 38; Conservative 22; Mismatches 70; Indels 22; Gaps 7;

QY 1 MCGGGS-----RADAIERYYESWTRTESTWLTYTDSALPSAATDSQPEAGL 51
 DB 860 LGAAGSSSLAPLVPAAPNPPEPTTSGSQRVENWEAH--SGLPFISSRSSP----L 913
 QY 52 HAGVLEDGLSSNGVLRPAAPPGIANPEKKNCG--GTQCPNSQNTSSGPL---TQKNGLM 106
 DB 914 QNLTLQEEHPAPSESADAVRGC-AGPDAGKHCVTGPSSRSRHSCTSGELATATVAHQSAA 972
 QY 107 ATEAKRDARMSAREVAIVNTEN-IRQMDRSK 137
 DB 973 ASGSASSTYFSSTDYASEVSENRRPDROR 1004

Search completed: April 12, 2005, 15:44:46
 Job time : 22.6667 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 15:16:23 ; Search time 94.6667 Seconds

(without alignment)
784.347 Million cell updates/sec

Title: US-10-705-716a-8

Perfect score: 767
Sequence: 1 MCGGSRRAAIERPRYESWT.....VTENIRQMDRSKVTNKCIN 145

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : uniprot_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	767	100.0	145	2	Q8VHV1 mus musculu
2	746	97.3	145	2	Q920K5
3	649	84.6	145	2	Q9HA93
4	634	82.7	145	2	Q8WNE9
5	621.5	81.0	180	2	Q8WXS3
6	480	62.6	123	2	Q9CYS9
7	472.5	61.6	149	2	Q8WXS1
8	294	38.3	54	2	Q8VBS8
9	294	38.3	54	2	Q790N3
10	288	37.5	73	2	Q8WXS0
11	278	36.2	54	2	Q8WTP6
12	272	35.5	80	2	Q8WXS2
13	267	34.8	54	2	Q8WNE8
14	161	21.0	32	2	Q801V5
15	87.5	11.4	1073	1	PVDA_PLAKN
16	86	11.2	1449	2	Q7XPB1
17	85	11.1	882	1	AREA ASPNG
18	84.5	11.0	367	2	Q7K2P9
19	84.5	11.0	494	2	Q8IR46
20	84.5	11.0	495	2	Q961T1
21	84.5	11.0	591	2	Q83H42
22	84.5	11.0	591	2	Q83NV7
23	84.5	11.0	659	2	Q9VV16
24	84.5	11.0	773	2	Q9FNV4
25	84	11.0	253	2	Q661S5
26	84	11.0	365	1	POLG_SUNVS
27	84	11.0	1433	1	CAT8_YEAST
28	83.5	10.9	291	2	Q871V5
29	83.5	10.9	865	1	NRFA_PENUR
30	83	10.8	476	2	Q7PV51
31	83	10.8	595	2	Q7UZ29

ALIGNMENTS

32	83	10.8	860	1	AREA PENRO	013508 penticillium
33	81.5	10.6	214	2	Q8MTG2	Q8mtc2 leucophaea
34	81.5	10.6	1100	2	Q91IA3	Q91ia3 white spot
35	81	10.6	219	2	Q82NV5	Q82nv5 streptomyc
36	81	10.6	228	2	Q91TQ3	Q91tq3 tupalid her
37	81	10.6	516	2	Q74641	Q74641 aspergillus
38	81	10.6	1070	1	PVDA_PLAKN	P50494 plasmodium
39	80.5	10.5	155	2	Q6U5F5	Q6u5f5 plasmodium
40	80.5	10.5	852	1	AOR2_HUMAN	O60341 homo sapien
41	80.5	10.5	2516	1	CCAD_DROME	Q24270 drosophila
42	80.5	10.5	2519	2	Q81P25	Q81p25 drosophila
43	80.5	10.5	2552	2	Q81P24	Q81p24 drosophila
44	80	10.4	124	2	Q868X0	Q868x0 plasmodium
45	80	10.4	516	1	INUZ_ASPII	Q94220 aspergillus

RESULT 1

Q8VHV1	PRELIMINARY;	PRT;	145 AA.
AC Q8VHV1			
DT 01-MAR-2002 (TREMBLrel. 20, Created)			
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)			
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)			
DE BAKC isoform 1-6-8 (Mus musculus 16 days neonate cerebellum cDNA, RIKEN full-length enriched library, clone:9630028H16 product:brain and DE acute leukemia, cytoplasmic, full insert sequence).			
GN Name=Bakc;			
OS Mus musculus (Mouse).			
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX NCBI_TaxID=10090;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC STRAIN=129S6/SVEVAC;			
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;			
RA Tanner S.M., Austin U.L., Leone G., Rush U.J., Plasse C., Heinonen K., Mrozek K., Sill H., Knuutila S., Kolitz J.E., Archer K.J.,			
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;			
RT "BAIC, the human member of a novel mammalian neuroectoderm gene lineage, is implicated in hematopoiesis and acute leukemia.";			
RT Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).			
RL [2]			
RP SEQUENCE FROM N.A.			
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;			
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;			
RA RIKEN FANTOM Consortium;			
RT "High-efficiency full-length cDNA cloning.";			
RT Meth. Enzymol. 303:19-44(1999).			
RL [3]			
RP SEQUENCE FROM N.A.			
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;			
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;			
RA RIKEN FANTOM Consortium;			
RT "Functional annotation of a full-length mouse cDNA collection.";			
RT Nature 409:685-690(2001).			
RL [4]			
RP SEQUENCE FROM N.A.			
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;			
RA The FANTOM Consortium;			
RT "The RIKEN Genome Exploration Research Group Phase I & II Team: Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";			
RT Nature 420:563-573(2002).			
RL [5]			
RP SEQUENCE FROM N.A.			
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;			
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;			
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;			
RA "Normalization and subtraction of cap-trapper-selected cDNAs to			

```

RT prepare full-length cDNA libraries for rapid discovery of new genes."
RL Genome Res. 10:1617-1630(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Kitajima T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishimura T., Harada A.,
RA Yamamoto R., Matsunoto H., Sakaguchi S., Ikegami T., Kaishima K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe K.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuda S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "Riken integrated sequence analysis (RISA) system-384-format
RL sequencing pipeline with 384 multicapillary sequencer."
RN Genome Res. 10:1757-1771(2000).
RP [7]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RA Aachai J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imocani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katon H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF371320; AAL50516.1; -.
DR EMBL; AK079337; BAC37611.1; -.
DR MGI; MGI:1928704; BaalC.
DR Pfam; PF06989; BAALC_N; 1.
SQ SEQUENCE 145 AA; 15515 MW; 4972670A618C4D6D CRC64;

Query Match 100.0%; Score 767; DB 2; Length 145;
Best local Similarity 100.0%; Pred. No. 5, 1e-65;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGGSRADAIIEPRYYESWTRETESTWLTYYTSDALPSAAATDSGPBAGGLHAGVLEDDG 60
DB 1 MCGGSRADAIIEPRYYESWTRETESTWLTYYTSDALPSAAATDSGPBAGGLHAGVLEDDG 60
QY 61 SSSNGVLRPAAPGCIANPEKKNCGTCCPSQNLSSGPLETKONGIMWATEKRDARMSAR 120
DB 61 SSSNGVLRPAAPGCIANPEKKNCGTCCPSQNLSSGPLETKONGIMWATEKRDARMSAR 120
QY 121 EVAINTENIRQMDRSKRYTKNCIN 145
DB 121 EVAINTENIRQMDRSKRYTKNCIN 145

RESULT 2
Q920K5 PRELIMINARY; PRT; 145 AA.
AC Q920K5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 05-JUL-2004 (TrEMBLrel. 19, last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, last annotation update)
DR Dem-A20-4 (BAA1C isoform 1-6-8).
GN Name=Dem-A20-4; Synonyms=BaalC;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang X., Tian Q., Li W., Okano A., Suzuki T.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=Sprague-Dawley;
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Anestin J.L., Leone G., Rush L.J., Plasch C., Heimonen K.,
RA Mrozek K., Sill H., Kruetli S., Kolitz J.E., Archer C.J.,
RA Calligaris M.A., Bloomfield C.D., de la Chapelle A.;
RT "BAA1C, the human member of a novel mammalian neurocorterm gene
RL lineage, is implicated in hematopoiesis and acute leukemia."
DR Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AB073319; BAB70507.1; -.
DR EMBL; AF371321; AAL50517.1; -.
DR Pfam; PF06989; BAALC_N; 1.
SQ SEQUENCE 145 AA; 15475 MW; D5A27AD67456F341 CRC64;

Query Match 97.3%; Score 746; DB 2; Length 145;
Best local Similarity 97.2%; Pred. No. 5, 1e-63;
Matches 141; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MCGGSRADAIIEPRYYESWTRETESTWLTYYTSDALPSAAATDSGPBAGGLHAGVLEDDG 60
DB 1 MCGGSRADAIIEPRYYESWTRETESTWLTYYTSDALPSAAATDSGPBAGGLHAGVLEDDG 60
QY 61 SSSNGVLRPAAPGCIANPEKKNCGTCCPSQNLSSGPLETKONGIMWATEKRDARMSAR 120
DB 61 SSSNGVLRPAAPGCIANPEKKNCGTCCPSQNLSSGPLETKONGIMWATEKRDARMSAR 120
QY 121 EVAINTENIRQMDRSKRYTKNCIN 145
DB 121 EVAINTENIRQMDRSKRYTKNCIN 145

RESULT 3
Q9HA93 PRELIMINARY; PRT; 145 AA.
AC Q9HA93;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, last annotation update)
DR Hypothetical protein FLJ12015 (BAA1C isoform 1-6-8) (Brain and acute
DE leukemia, cytoplasmic) (BAA1C 1-6-8).
GN Name=BAA1C;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole embryo;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Oobayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Shimizu T., Nomura Y.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamihara K., Katsura N., Sato K., Tanikawa M., Yamazaki M.,
RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hiraoa S., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
RA Togoya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Musashino K., Yuki H., Oshima A., Sasaki N., Aotaka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Teraishi Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiroaki M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,

```

RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuno Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isegai T., Sugano S.;
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs.";
 RL Nat. Genet. 36:40-45(2004).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
 RA Tamer S.M., Auestin J.L., Leone G., Rush L.J., Plasch C., Heimonen K.,
 RA Mrozek K., Sill H., Knutti S., Kolitz J.E., Archer K.J.,
 RA Calliguri M.A., Bloomfield C.D., de la Chapelle A.;
 RT "BALC, the human member of a novel mammalian neuroectoderm gene
 RT lineage, is implicated in hematopoiesis and acute leukemia.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uadin T.B., Toshyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mallary S.J.,
 RA Bosak S.A., Mcwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultky S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Ketterman M., Madan A.C., Rodchenko Y., Sanchez A.,
 RA Whiting M., Madan A.C., Shevchenko V., Bouffard G.C.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smalusz D.E., Schermer A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Strausberg R.;
 RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK022077; BAB13960.1; -;
 DR EMBL; AF371319; AAL50515.1; -;
 DR EMBL; BC011517; AAL1517.1; -;
 DR EMBL; AF363578; AAL50377.1; -;
 DR InterPro; IPR009728; BAALC_N;
 DR Pfam; PF06989; BAALC_N; 1.
 SQ SEQUENCE 145 AA; 15551 MW; CF892BBE283DD92E CRC64;

Query Match 84.6%; Score 649; DB 2; Length 145;
 Best Local Similarity 83.4%; Pred. No. 8.5e-54;
 Matches 121; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

QY 1 MCCGSRADAIIPRYESWTRETESTWLTYSDALPSAAATDSCPEAGGLHAGVLEDEL 60
 DB 1 MCCGSRADAIIPRYESWTRETESTWLTYSDALPSAAATDSCPEAGGLHAGVLEDEL 60
 QY 61 SNGVLPRAAPGGINPEKKMCGTCPPSONLSSGPLETKOKNGLWATEAKRDARMSAR 120
 DB 61 PSNGVPRSTAPGGINPEKKMCGTCPPSONLSSGPLETKOKNGLWATEAKRDARMSAR 120
 QY 121 EVAIVNTENIRQMDRSKRVTKNCIN 145
 DB 121 EVTIVNTDSIQMDRSKRVTKNCIN 145

RESULT 4
 Q8WNE9 PRELIMINARY; PRT; 145 AA.
 AC Q8WNE9;

DT 01-MAR-2002 (TREMblrel. 20, Created)
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
 DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)
 DE BAALC isoform 1-6-8.
 GN Name=BAALC;
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
 RA Tamer S.M., Auestin J.L., Leone G., Rush L.J., Plasch C., Heimonen K.,
 RA Mrozek K., Sill H., Knutti S., Kolitz J.E., Archer K.J.,
 RA Calliguri M.A., Bloomfield C.D., de la Chapelle A.;
 RT "BALC, the human member of a novel mammalian neuroectoderm gene
 RT lineage, is implicated in hematopoiesis and acute leukemia.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
 DR EMBL; AF371322; AAL50518.1; -;
 DR Pfam; PF06989; BAALC_N; 1.
 SQ SEQUENCE 145 AA; 15401 MW; C75BD7D0E8F82E26 CRC64;

Query Match 82.7%; Score 634; DB 2; Length 145;
 Best Local Similarity 82.1%; Pred. No. 2.3e-52;
 Matches 119; Conservative 11; Mismatches 15; Indels 0; Gaps 0;

QY 1 MCCGSRADAIIPRYESWTRETESTWLTYSDALPSAAATDSCPEAGGLHAGVLEDEL 60
 DB 1 MCCGSRADAIIPRYESWTRETESTWLTYSDALPSAAATDSCPEAGGLHAGVLEDEL 60
 QY 61 SNGVLPRAAPGGINPEKKMCGTCPPSONLSSGPLETKOKNGLWATEAKRDARMSAR 120
 DB 61 PSNGVPRSTAPGGINPEKKMCGTCPPSONLSSGPLETKOKNGLWATEAKRDARMSAR 120
 QY 121 EVAIVNTENIRQMDRSKRVTKNCIN 145
 DB 121 EVTIVNTDSIQMDRSKRVTKNCIN 145

RESULT 5
 Q8WKS3 PRELIMINARY; PRT; 180 AA.
 AC Q8WKS3;
 DT 01-MAR-2002 (TREMblrel. 20, Created)
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
 DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)
 DE BAALC 1-5-6-8.
 GN Name=BAALC;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
 RA Tamer S.M., Auestin J.L., Leone G., Rush L.J., Plasch C., Heimonen K.,
 RA Mrozek K., Sill H., Knutti S., Kolitz J.E., Archer K.J.,
 RA Calliguri M.A., Bloomfield C.D., de la Chapelle A.;
 RT "BALC, the human member of a novel mammalian neuroectoderm gene
 RT lineage, is implicated in hematopoiesis and acute leukemia.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
 DR EMBL; AF363578; AAL50379.1; -;
 DR Pfam; PF06989; BAALC_N; 1.
 SQ SEQUENCE 180 AA; 19224 MW; 380183E0F188F684 CRC64;

Query Match 81.0%; Score 621.5; DB 2; Length 180;
 Best Local Similarity 67.2%; Pred. No. 4.5e-51;
 Matches 121; Conservative 10; Mismatches 14; Indels 35; Gaps 1;

QY 1 MCCGSRADAIIPRYESWTRETESTWLTYSDALPSAAATDSCPEAGGLHAGVLEDEL 53
 DB 1 MCCGSRADAIIPRYESWTRETESTWLTYSDALPSAAATDSCPEAGGLHAGVLEDEL 53

```

QY 54 -----GVLDDGLSSNGVLRPAAPGGIANPEKXNCGT 85
DB 61 KIKAPTVSDSEGLFASAKMPLAVFSGHGLDGLPSNGVPRSTAPGGIPNPEKXNCGT 120
QY 86 QCPNSQNTSSGPTLQKONGLMATEAKRDPAKMSAREVAINTENIRQMDRSKRVTKNCIN 145
DB 121 QCPNPSLSSGPTLQKONGLMATEAKRDPAKMSAREVAINTENIRQMDRSKRVTKNCIN 180

RESULT 6
Q9CYS9 PRELIMINARY; PRT; 123 AA.
AC Q9CYS9
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length
DE enriched library, clone:2810457D07 product:brain and acute leukemia,
DE cytoplasmic, full insert sequence.
GN Name=Baalc;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN PANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RT The PANTOM Consortium;
RA "The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayata N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=20330913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Saeki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Kitsuiri T., Taahiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Wachihi M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multiplexed capillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,

```

```

RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuo M.,
RA Hanganaki T., Hara A., Hayatsu N., Hiramoto K., Hirooka T., Hori F.,
RA Imomani K., Ishii Y., Itoh M., Izawa M., Kaakawa T., Kato H.,
RA Kawai J., Kojima Y., Kono H., Kouda M., Koy S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Teijima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK013358; BAB28808.1; -.
DR MGI; MGI:1928704; Baalc.
SQ SEQUENCE 123 AA; 13391 MW; 33DFCB5F8CFAA0AB CRC64;

Query Match
Best Local Similarity 98.6%; Score 480; DB 2; Length 123;
Matches 92; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 53 AGVLDDGLSSNGVLRPAAPGGIANPEKXNCGTQCPNSQNTSSGPTLQKONGLMATEAKR 112
DB 31 SGVLDDGLSSNGVLRPAAPGGIANPEKXNCGTQCPNSQNTSSGPTLQKONGLMATEAKR 90
QY 113 DAKRMSAREVAINTENIRQMDRSKRVTKNCIN 145
DB 91 DAKRMSAREVAINTENIRQMDRSKRVTKNCIN 123

RESULT 7
Q8WXSL PRELIMINARY; PRT; 149 AA.
AC Q8WXSL
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE BAALC 1-5-6-7-8.
GN Name=BAALC;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plass C., Heinonen K.,
RA Mrozek K., Sill H., Knuttila S., Kolitz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BAALC, the human member of a novel mammalian neuroectoderm gene
RT lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AF163578; AAL50381.1; -.
DR Genew; HGNC:14333; BAALC.
DR Pfam; PF06989; BAALC_N; 1.
SQ SEQUENCE 149 AA; 15577 MW; DB26C19669B91F6E CRC64;

Query Match
Best Local Similarity 61.6%; Score 472.5; DB 2; Length 149;
Matches 94; Conservative 3; Mismatches 12; Indels 35; Gaps 1;

```

```

QY 1 MCGGSRDAIEPRYESTRTSTWLTJYDSDALPSAAATDSGPEAGLHA----- 53
DB 1 MCGGSRDAIEPRYESTRTSTWLTJYDSDALPSAAATDSGPEAGLHVLEAKS 60
QY 54 -----GVLDDGLSSNGVLRPAAPGGIANPEKXNCGT 85
DB 61 KIKAPTVSDSEGLFASAKMPLAVFSGHGLDGLPSNGVPRSTAPGGIPNPEKXNCGT 120
QY 86 QCPNSQNTSSGPTLQKONGLMATE 109
DB 121 QCPNPSLSSGPTLQKONGLOTTE 144

RESULT 8

```


08VB88
ID 08VB88 PRELIMINARY; PRT; 54 AA.
AC 08VB88;
DT 01-MAR-2002 (TEMBLrel. 20, Created)
DT 01-MAR-2002 (TEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE BAALC isoform 1-8.
GN Name=BAALC;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=12986/SEVETAC;
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plase C., Heinson K.,
RA Mrozek K., Sill H., Knuutila S., Kolitz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BAALC, the human member of a novel mammalian neuroectoderm gene
lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AF371324; AAL50520.1; -
DR MGD; MGI:1928704; Baalc.
DR Pfam; PF06989; BAALC_N; 1.
SQ SEQUENCE 54 AA; 5667 MW; FEE8C0BBFBCDB829 CRC64;

Query Match 38.3%; Score 294; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 1.5e-20;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MCGGSRADAIERYESWTRETESTWLTYYTSDALPSAATDGPAGGLHAG 54
DB 1 MCGGSRADAIERYESWTRETESTWLTYYTSDALPSAATDGPAGGLHAG 54

RESULT 9
ID 0790N3 PRELIMINARY; PRT; 54 AA.
AC 0790N3;
DT 05-JUL-2004 (TEMBLrel. 27, Created)
DT 05-JUL-2004 (TEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TEMBLrel. 27, Last annotation update)
DE BAALC isoform 1-8.
GN Name=BAALC;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plase C., Heinson K.,
RA Mrozek K., Sill H., Knuutila S., Kolitz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BAALC, the human member of a novel mammalian neuroectoderm gene
lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AF371325; AAL50521.1; -
DR InterPro; IPR0069728; BAALC_N.
DR Pfam; PF06989; BAALC_N; 1.
SQ SEQUENCE 54 AA; 5667 MW; FEE8C0BBFBCDB829 CRC64;

Query Match 38.3%; Score 294; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 1.5e-20;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MCGGSRADAIERYESWTRETESTWLTYYTSDALPSAATDGPAGGLHAG 54
DB 1 MCGGSRADAIERYESWTRETESTWLTYYTSDALPSAATDGPAGGLHAG 54

RESULT 10
ID 08WXSO PRELIMINARY; PRT; 73 AA.
AC 08WXSO;
DT 01-MAR-2002 (TEMBLrel. 20, Created)
DT 01-MAR-2002 (TEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE BAALC 1-4-5-6-8.
GN Name=BAALC;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plase C., Heinson K.,
RA Mrozek K., Sill H., Knuutila S., Kolitz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BAALC, the human member of a novel mammalian neuroectoderm gene
lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AF363578; AAL50380.1; -
DR Pfam; PF06989; BAALC_N; 1.
SQ SEQUENCE 73 AA; 7671 MW; 98DBC2E6E6EF524A CRC64;

Query Match 37.5%; Score 288; DB 2; Length 73;
Best Local Similarity 91.4%; Pred. No. 7.8e-20;
Matches 53; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 MCGGSRADAIERYESWTRETESTWLTYYTSDALPSAATDGPAGGLHAG 58
DB 1 MCGGSRADAIERYESWTRETESTWLTYYTSDALPSAATDGPAGGLHAG 58

RESULT 11
ID 08WTP6 PRELIMINARY; PRT; 54 AA.
AC 08WTP6;
DT 01-MAR-2002 (TEMBLrel. 20, Created)
DT 01-MAR-2002 (TEMBLrel. 20, Last sequence update)
DT 05-JUL-2004 (TEMBLrel. 27, Last annotation update)
DE BAALC isoform 1-8 (BAALC protein) (BAALC 1-8).
GN Name=BAALC;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plase C., Heinson K.,
RA Mrozek K., Sill H., Knuutila S., Kolitz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BAALC, the human member of a novel mammalian neuroectoderm gene
lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins J.L., Wagner L., Scheffen C.M., Schuler G.D.,
RA Altshuler S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.U., Udels T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S.N., Garcia A.M., Gay L.J., Hultik S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Gilmwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skaleja U., Smalins D.E., Schermer A., Schein J.E.,
RA Jones S.J., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.,
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF371323; AAL50519.1; -
DR EMBL; BC035038; AAL50503.1; -
DR EMBL; AF363578; AAL50378.1; -
DR Pfam; PF06989; BAALC_N; 1.
SQ SEQUENCE 54 AA; 5663 MW; FEF084EABED9B829 CRC64;

Query Match 36.2%; Score 278; DB 2; Length 54;
Best Local Similarity 94.4%; Pred. No. 4.9e-19;
Matches 51; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERYESWTRETSWTLYTSDALPSAAATDGSPEAGGLHAG 54
Db 1 MCGGSRADAIERYESWTRETSWTLYTSDALPSAAATDGSPEAGGLHAG 54

RESULT 12
Q8WXS2 PRELIMINARY; PRT; 80 AA.
AC Q8WXS2;
DT 01-JUN-2003 (TREMBlrel. 20, Created)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DN BAALC 1-2.
GN Name=BAALC;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plas C., Heimonen K.,
RA Mrozek K., Sill H., Knutti S., Koltz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.,
RT "BAALC, the human member of a novel mammalian neuroectoderm gene
RT lineage, is implicated in hematopoiesis and acute leukemia."
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AF363578; AAL50382.1; -
DR Pfam; PF06989; BAALC_N; 1.
SQ SEQUENCE 80 AA; 8573 MW; 70BBE9B91C3245E2D CRC64;

Query Match 35.5%; Score 272; DB 2; Length 80;
Best Local Similarity 94.3%; Pred. No. 2.9e-18;
Matches 50; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERYESWTRETSWTLYTSDALPSAAATDGSPEAGGLHAG 53
Db 1 MCGGSRADAIERYESWTRETSWTLYTSDALPSAAATDGSPEAGGLHAG 53

RESULT 13
Q8WNE8 PRELIMINARY; PRT; 54 AA.
AC Q8WNE8;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DE 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DN BAALC isoform 1-8.
GN Name=BAALC;
OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plas C., Heimonen K.,
RA Mrozek K., Sill H., Knutti S., Koltz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.,
RT "BAALC, the human member of a novel mammalian neuroectoderm gene
RT lineage, is implicated in hematopoiesis and acute leukemia."
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AF371326; AAL50522.1; -
DR Pfam; PF06989; BAALC_N; 1.
SQ SEQUENCE 54 AA; 5681 MW; 6538C3DABED9B825 CRC64;

Query Match 34.8%; Score 267; DB 2; Length 54;
Best Local Similarity 92.6%; Pred. No. 5.4e-18;
Matches 50; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERYESWTRETSWTLYTSDALPSAAATDGSPEAGGLHAG 54
Db 1 MCGGSRADAIERYESWTRETSWTLYTSDALPSAAATDGSPEAGGLHAG 54

RESULT 14
Q801V5 PRELIMINARY; PRT; 32 AA.
AC Q801V5;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DE 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE SI:2C215113.3 (Novel protein similar to human brain and acute
DE leukemia, cytoplasmic (BAALC)) (Fragment).
GN Name=SI:2C215113.3;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Sehra H.,
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL807244; CAD87801.1; -
DR InterPro; IPR009728; BAALC_N.
DR Pfam; PF06989; BAALC_N; 1.
FT NON_TER 32
FT 32
SQ SEQUENCE 32 AA; 3663 MW; 916A5445D263B7F7 CRC64;

Query Match 21.0%; Score 161; DB 2; Length 32;
Best Local Similarity 90.6%; Pred. No. 3.4e-08;
Matches 29; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERYESWTRETSWTLYTSD 32
Db 1 MCGGSRADAIERYESWTRETSWTLYTSD 32

RESULT 15
P22545 STANDARD; PRT; 1073 AA.
ID P22545;
AC P22545;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 29-MAR-2004 (Rel. 43, Last annotation update)
DE Duffy receptor, alpha form precursor (Erythrocyte binding protein).
OS Plasmodium knowlesi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5850;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92357776; PubMed=1496004;

RA Adams J.H., Sim B.K., Dolan S.A., Fang X., Kaslow D.C., Miller L.H.;
 RT "A family of erythrocyte binding proteins of malaria parasites.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:7085-7089(1992).
 RN [2]
 RP SEQUENCE OF 296-1073 FROM N.A.
 RX MEDLINE=91004213; PubMed=2170017; DOI=10.1016/0092-8674(90)90295-P;
 RA Adams J.H., Hudson D.E., Toril M., Ward G.E., Wellens T.E., Akawa M.,
 Miller L.H.;
 RT "The Duffy receptor family of Plasmodium knowlesi is located within
 the micronemes of invasive malaria merozoites.";
 RL Cell 63:141-153(1990).
 CC -1- FUNCTION: Binds to the human erythrocytes Duffy blood group
 determinant.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M90466; AAA29602.1; -;
 DR EMBL; M68517; AAA29590.1; -;
 DR EMBL; M68518; AAA29591.1; -;
 DR PIR; A35970; A35970.
 DR InterPro; IPR008602; Duffy_binding.
 DR Pfam; PF05424; Duffy_binding; 1.
 KW Glycoprotein; Malaria; Multigene family; Receptor; Signal;
 KW Transmembrane.
 FT SIGNAL 1 21 Potential.
 FT CHAIN 22 1073 Duffy receptor, alpha form.
 FT DOMAIN 22 1007 Extracellular (Potential).
 FT TRANSMEM 1008 1029 Potential.
 FT DOMAIN 1030 1073 Cytoplasmic (Potential).
 FT CARBOHYD 134 134 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 179 179 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 202 202 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 252 252 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 348 348 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 679 679 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 746 746 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 779 779 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 788 788 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 1073 AA; 120683 MW; 3965FC9F46B71808 CRC64;

Query Match 11.4%; Score 87.5; DB 1; Length 1073;
 Best Local Similarity 31.8%; Pred. No. 24;
 Matches 28; Conservative 13; Mismatches 44; Indels 3; Gaps 3;

QY 31 TSDALPSAATDGPAGAGLAVLEGLSSNGV-LRPAPGIANPEKXNC-GTQCP 88
 Db 647 TVSSDVPVSGKSDSPSTASHALAGENGHNHTDTEPKEDKADPQKDIKVKQKDT 706

QY 89 NSGNLSS-GPLTRKQNGWATEAKRPAK 115
 Db 707 DDRSGSLGPHHTDRAATLGRTHEKQTE 734

Search completed: April 12, 2005, 15:42:12
 Job time : 96.6667 secs

This Page Blank (uspto)